Morphological Variability in Second Language Spanish

Corrine Lee McCarthy

Department of Linguistics

McGill University, Montreal

April, 2007

A thesis submitted to McGill University

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-32312-0
Our file Notre référence
ISBN: 978-0-494-32312-0

ABSTRACT

Research on morphological variability in second language (L2) acquisition has focused on the syntactic consequences of variability: that is, whether or not morphological variability entails underlying syntactic deficits. The interrelationship between morphological features in their own right has been largely ignored. This thesis addresses the representation of L2 features by investigating the use of default morphology—the outcome of systematic substitution errors employed by speakers of L2 Spanish. It is hypothesized that underspecified features act as defaults; by assumption, those features that are unmarked are underspecified.

Evidence to support this hypothesis comes from two sets of experiments conducted on intermediate- and advanced-proficiency L2 Spanish subjects (L1 English). The first set of experiments addresses verbal morphology, and consists of a spontaneous production experiment on person, number, tense, and finiteness, and a comprehension task on person and number. The second set of experiments addresses gender and number in nominal morphology, and consists of a spontaneous production experiment on determiners, an elicited production experiment on clitics and adjectives, and a picture-selection task on the comprehension of clitics. Across tasks and across verbal and nominal domains, errors involve the systematic substitution of underspecified morphology. The observation that morphological variability extends to comprehension, and is qualitatively similar to the variability found in production, counters the suggestion that variability is strictly a product of mere performance limitations on production. Finally, the systematicity of substitution errors suggests that the natural classes of features such as gender, number, tense, and person are acquirable in an L2, regardless of whether or not these features have been instantiated in the native language.

RÉSUMÉ

La plupart des travaux de recherche portant sur la variabilité morphologique dans l'acquisition d'une langue seconde ont examiné les conséquences syntaxiques de cette variabilité et si, plus précisément, cette variabilité reflète une déficience syntaxique de base. La relation directe entre les différents traits morphologiques a été largement ignorée. La présente recherche vise à examiner la représentation des traits morphologiques chez les apprenants d'une langue seconde (dans ce cas, l'espagnol), en étudiant leur utilisation de la 'morphologie par défaut', c'est-à-dire, les erreurs morphologiques qu'ils produisent systématiquement. L'objectif des expériences décrites dans ce document est de tester l'hypothèse voulant que les traits qui sont sous-specifiés sont produits par défaut (notez qu'on présume que les traits non-marqués sont sous-spécifiés).

Afin d'appuyer cette hypothèse, on présente les résultats de deux séries d'expériences menées auprès d'apprenants de langue maternelle anglaise ayant une maîtrise intermédiaire à avancé de l'espagnol. La première série d'expériences visait la morphologie verbale, notamment la personne, le nombre, le temps et la distinction fini-infini dans le contexte d'une tâche de production de langage spontanée, ainsi qu'une tâche de compréhension qui examinait la personne et le nombre. La deuxième série d'expériences examinait la production du genre et du nombre des noms, dans le contexte d'une tâche de production spontanée visant les déterminants, une tâche de production suscitée visant les clitiques et les adjectifs, et une tâche où les participants devaient choisir une image, qui visait la compréhension des clitiques. Pour toutes les tâches expérimentales, et pour la morphologie verbale ainsi que nominale, les erreurs reflétaient la substitution systématique de traits morphologiques sous-spécifiés. Le constat que la variabilité morphologique est présente en compréhension et ressemble, de façon qualitative, à la variabilité identifiée en production va à l'encontre de la notion que la variabilité n'est qu'une conséquence indirecte des limites de la performance. Enfin, la nature systématique des erreurs de substitution observées semble indiquer que les apprenants de langues secondes sont capables d'acquérir des classes naturelles de traits morphologiques, telles que le genre, le nombre, le temps et la personne, même si ces traits ne sont pas présents dans la grammaire de leur langue maternelle.

ACKNOWLEDGEMENTS

First of all, I owe many thanks to my dear friend and colleague Elena Valenzuela. Without Elena, this dissertation would have taken a lot longer than it did. Elena granted me access to a large group of L2 Spanish speakers. She also connected me to Ana Faure, who graciously agreed to run experiments on my behalf; I owe her many thanks for her careful data collection. I am also thankful to the native Spanish speakers who have assisted me in data collection and granted me judgment after judgment, particularly Eva Villalba.

I would like to acknowledge some of my McGill student colleagues who lent their support during the writing of this thesis: Deena Fogle, Lotus Goldberg, Jen Mah, Jason McDevitt, Kristi Supinski, and Mari Umeda. Deena and Jen, along with Eric Farquhar, were especially helpful in serving as models for my picture description task—they put on earrings, sold books, ate apples, and washed dishes with great conviction.

I also wish to acknowledge my parents for supporting me during the final stages of this dissertation, and my cats Hank and Queenie for making working at home a pleasure. My Chicago colleagues Mike and Cely deserve an acknowledgement for their kind and calming words of encouragement (served daily with peanut butter).

In terms of financial support, I also owe thanks to McGill University and the Faculty of Arts for their financial support in the form of graduate student fellowships, travel grants, and fee waivers; to the J. W. McConnell foundation for 3 years of support in the form of a McGill Major Fellowship; and to the Center for Research on Language, Mind and Brain for funding to travel and disseminate this research. This research was also funded by SSHRCC grants 410-2001-0719 and 410-2004-1774 (to Lydia White), and FQRSC grants 2001-ER-66973 and 2006-SE-103710 (to Lydia White et al.).

In addition, I would like to acknowledge a few researchers whose helpful comments, made over the course of presenting and/or discussing my dissertation research, eventually were incorporated into the final product in one way or another: Richard Cameron, Roumyana Slabakova, and Virginia Valian.

Finally, a super huge amount of gratitude goes to the dissertation committee and defense committee members. Andrea Gualmini proved to be a great resource for

experimental tasks (and espresso). I am grateful for his keen eye to detail and amazingly insightful feedback (which was sometimes provided within minutes. Susana Bejar has given me a lot to think about in the future; I am also grateful that she reminded me of the importance of clear generalizations and interesting data—the things that persist even when theories evolve. I would like to acknowledge Joyce Bruhn de Garavito and Roumyana Slabakova for their careful comments on an earlier version of this thesis. I am particularly grateful to Joyce for her sharp insights into Spanish morphosyntax, and her kind words of encouragement. Glyne Piggott and Charles Boberg brought insightful questions to my dissertation defense, and the answers to these questions will undoubtedly strengthen my future research.

I would also like to thank the following linguists, all of whom have been influential in my development over the years: Charles Boberg, Jonathan Bobaljik, Brendan Gillon, Glyne Piggott, Susi Wurmbrand, Heather Goad, Alan Juffs at the University of Pittsburgh. And, most importantly, I would like to acknowledge Lydia White, who supervised this dissertation: thank you for providing superb examples of clear argumentation, for offering speedy and insightful feedback, for forcing me to elaborate when I thought I'd already made my point (when of course I hadn't!), and for believing that the questions that I wanted to address were, in fact, worth addressing.

TABLE OF CONTENTS

List of abbreviations

1	1st person
2	2nd person
3	3rd person
AA	attributive adjective
ACC	accusative (case)
AGR(P)	agreement (phrase)
C(P)	complementizer (phrase)
CL	clitic
COM	common gender
DAT	dative (case)
DO	direct object
D(P)	determiner (phrase)
f(em)	feminine
FFFH	Failed Functional Features hypothesis
FTFA	Full Transfer Full Access hypothesis
inf	infinitive
IO	indirect object
I(P)	inflectional (phrase)
L1	first language
L2	second language
L2er	second language learner
m(asc)	masculine
MSIH	Missing Surface Inflection Hypothesis

MUSH	Morphological Underspecification Hypothesis
neg	negation
NEUT	neuter gender
N(P)	noun (phrase)
NOM	nominative (case)
PA	predicative adjective
PF	phonetic form
pl	plural
POSS	possessive
pres	present tense
s(g)	singular
T(P)	tense (phrase)
UG	Universal Grammar
V(P)	verb (phrase)

Chapter 1
Introduction

This dissertation seeks to understand a set of rules and representations that govern the second-language (L2) grammar—in particular, those rules and representations that pertain to inflectional morphology. I will treat the L2 system, often called an *interlanguage*, as a grammar: a productive system that is rule-governed and that constitutes the linguistic *competence* (or abstract knowledge, in the sense of Chomsky 1965) of the L2 learner (L2er). A controversial assumption in L2 theory concerns the availability of Universal Grammar (UG): are these L2/interlanguage grammars constrained by the same universal principles that constrain native grammars? Though the main goal of this dissertation is not to argue that UG is available in L2 acquisition, I will operate under the assumption that what the learner is building is, in fact, a grammar, and therefore is an example of a human language. This dissertation is therefore devoted to establishing a particular set of the representations that underlie the linguistic competence of the users of such grammars; in particular, those representations that apply to the domain of inflectional morphology.

This dissertation is carried out within a generative framework, which holds that the underlying representations are manipulated by a computation system to generate what we observe as language. A central issue in generative L2 theory concerns the extent to which acquisition of a native or first language (L1) is similar to acquisition of an L2. While the outcome of L1 acquisition is never in doubt, adult L2 acquisition is subject to all sorts of errors and inconsistencies. One area where persistent errors and

inconsistencies arise is in tense and agreement. The inconsistent use of target-like morphology is often termed *morphological variability.*

The existence of morphological variability is surprising in a sense: there are abundant examples of inflection in the input, yet L2ers do not seem able to make use of such information. For example, in the L2 classroom or in a place where a L2 is spoken, we would expect the learner to hear plenty 1st person and past-tense verbs, and feminine and plural clitics; however, L2ers frequently fail to accurately produce (and, as I will show here, comprehend) the inflectional morphology that corresponds to these forms. As previous research (e.g. Lardiere 1998a,b, White et al 2004, and many others) has shown and as I will show here, these problems in the comprehension and production of inflectional morphology are often highly persistent, even for learners at very high levels of proficiency.

Theories disagree over what morphological variability means. On the one hand, variability may be symptomatic of an underlying syntactic deficit (e.g. Clahsen 1988, Hawkins & Chan 1997, and many others), or they may arise from something more superficial (e.g. Prévost and White 2000b, Montrul 2004, and many others). Questions arise for both perspectives on variability. If, under the first perspective, syntax is somehow impaired, how do L2ers manage to sometimes get it right? If, under the second perspective, the underlying syntactic representations are "there", why should variability arise at all?

While the debate over the syntactic consequences of morphological variability continues, I will focus my attention on another question that has been largely ignored: given the undeniable existence of variability, how can a theory constrain its outcome?

As I stated previously, one approach holds that variability results from mere performance issues (in the sense of Chomsky 1965). Yet if many of the errors that we consider to be 'performance' are actually systematic rather than random, as I will attempt to show, we are justified in considering whether they might be (partially) governed by linguistic competence.[1]

Decades of L2 research tell us that variability does not mean "anything goes". Errors are systematic, suggesting that underlying rules or representations are responsible for their systematic character. Yet the properties of these underlying representations are far from clear. As Archibald notes in the introductory linguistics textbook *Contemporary Linguistic Analysis*, "L2 morphology has been studied more or less in a theoretical vacuum" (Archibald 2004: 360). The major goal of this dissertation is to try to fill this void by establishing some principles that 1) govern the representation of L2 morphological features, 2) explain why variability is systematic, and 3) explain why variability favors the use of one variant (the "default") over another.

As a starting point, let us consider a simple set of data. The following (hypothetical) sentences, all of which are ungrammatical according to the grammar of a native speaker (NS), are typical of the kind of errors L2ers make.

[1] This statement follows the logic of sociolinguistic theory (e.g. Labov 1972): the systematic correlation between social factors and linguistic factors suggests that variation is rule-governed and therefore part of linguistic competence. The systematicity of variation means either that variation should not be dismissed as mere performance and instead included in a theory of competence, or the strict separation between competence and performance should be discarded. In this dissertation, I look at language-internal factors that contribute to variability, as opposed to language-external, social factors; a central goal is to account for as much of this systematic variation as possible by appealing to the representation of features.

Plausible examples:

1. a) Ellos habla

They speak.SG

b) Ayer yo camino

Yesterday I walk-1SG

c) camisa blanco

shirt-FEM white-MASC

(1a) involves the substitution of a singular verb (*habla)* for a plural one (*hablan).* (1b) involves the substitution of a present-tense verb (*camino)* for a past-tense one (*caminé* '(I) walked'). (1c) involves the substitution of a masculine adjective (*blanco*) for a feminine one (*blanca* 'white-FEM').

Rare examples:

2. a) Él hablan

He speak-PL

b) Ahora yo caminé (intended meaning: present-tense)

Now I walked-1SG

c) zapato blanca

shoe-MASC white-FEM

(2a) involves the substitution of a plural verb in a context where a singular verb should have occurred (*habla* 'speak.3SG'). (2b) involves the substitution of a past-

tense verb in a context where a present-tense one should have occurred (*camino* '(I) walk'). (2c) involves the substitution of a feminine adjective for a masculine one.

What accounts for the differences between (1) and (2)? In this dissertation, I will argue that the explanation lies in *morphological features* and how they are represented in the L2 grammar. Morphological features are abstract units of grammatical information that encode masculine, feminine, present, past, singular, plural, and so on. An overarching goal of this dissertation is to attempt to formulate a theory of how features are represented in the L2 grammar; this theory will attempt to explain systematic patterns of variability like the ones we have just seen in (1) and (2).

Chapter Two introduces the key notions involved in this discussion: morphological features, their representations, and how these representations interact with the syntactic component of the grammar. In this chapter, I introduce *underspecification* and *markedness*, concepts I will exploit in order to capture generalizations on error types. I elaborate on the necessary assumptions for the grammar, given the adoption of underspecification.

Chapter Three outlines the major generalizations about morphological variability, with specific reference to L2 Spanish, where relevant data exist. I also discuss how variability has been explained and accounted for in L2 theory. Finally, I state the predictions of the Morphological Underspecification Hypothesis (MUSH) for L2

acquisition, which I propose as an account for morphological variability, and compare its predictions to those made by some alternative accounts.

Chapters Four and Five consist of experimental tests of the hypothesis, in production and comprehension, across domains. Chapter Four looks at morphological variability in the verbal domain; specifically, I consider variability in person, number, tense (past versus present), and finiteness. Chapter Five considers variability in the nominal domain; specifically, gender and number agreement in determiners, adjectives, and direct-object (DO) clitics.

Chapter Six summarizes the major findings of this dissertation, and suggests new directions for future research.

Chapter 2
Features, Markedness, and the Syntax-Morphology Interface

This chapter gives a theoretical review of some of the issues concerning features, their representation, and their interface with syntax. Nearly all of the generative L2 acquisition literature has assumed, either tacitly or explicitly, that features are unstructured bundles lacking any sort of organizing principle (see, for example, Lardiere 2005, Montrul 2004, and many others). Yet recent work in theoretical morphology demonstrates that features show organization and structure; one proposal holds that *underspecification*, the elimination of unnecessary or redundant information from a representation, applies to morphological features.[2] In this chapter, I will argue that underspecification provides the basis for an explanatory account of morphological variability in L2 acquisition. In L1 acquisition, evidence supports the claim that there is organization and structure to morphological features; for instance, Harley and Ritter (2002) rephrase existing literature on L1 acquisition in terms of a feature geometric approach to the organization of features, and find some support for the universality of their claims; specifically, data suggest that children acquire the least specified forms prior to those that are more specified, and correspondingly bear more structure. To my knowledge, their claims have not been applied to L2 acquisition.

A few studies (e.g. Prévost and White 2000b) argue for underspecification—and therefore organization and structure—based on what defaults surface, but this argument lacks independent evidence for default feature representations and hence suffers from circularity. Lardiere (2005) argues for the study of feature assembly and organization in addressing the problem of variability, but to this point, the specific set of features that are problematic is still largely unknown. Thus, for the most part, the study of the organization of L2 features is still in its infancy.

[2] This follows in the tradition of phonology. The use of feature geometries to encode underspecification has been influential in phonology since the 1980s (e.g. Clements 1985, Sagey 1986) and feature geometries are now influential morphology as well (e.g. Harley 1994, Bonet 1995).

The goal of this chapter will be to motivate a set of assumptions regarding how morphological features are represented, relying heavily on the notion of underspecification. The proposed representations will be tested in the chapters that follow.

Before beginning the discussion of features and their representations, I will define some of the terms I will use throughout the remainder of this thesis. These definitions follow largely from White (2003). I take a *functional feature* to be an abstract unit expressing a grammatical property like [past], [feminine], or [plural]; these are notated in square brackets. I will frequently refer to these as simply *features*. I will also refer to the larger (in the sense that they may encompass other features) groupings of tense, gender, and number as features. *Functional categories*, on the other hand, correspond to syntactic units like *determiner*, *Infl*, and *Comp*, and are associated with functional features. For example, the functional category *determiner* is associated with the functional features gender and number in Spanish.

Functional features like the ones identified above correspond to morpho-phonological strings, or morphemes, in ways that are sometimes straightforward, and sometimes complex. (By using the term *morpheme* to mean the phonological exponent of morphology, I follow the acquisition literature; in the theoretical literature on morphology, terminology is different, as we will see during the elaboration of Distributed Morphology in Section 2.4.) One morpheme can encode exactly one feature, as in the *–s* of *casas* 'houses', in which *–s* corresponds to [plural] (cf. *casa* 'house). However, there are many examples in which a single morpheme corresponds to more than one feature. For example, *–é* in *hablé* 'I spoke' corresponds to [1], [singular], and [past]. The complex interrelationship between features and overt realizations requires that a number of theoretical assumptions be made in order to correctly account for natural language data. I will discuss some of these assumptions in detail.

This chapter is organized as follows. In Section 2.1, I describe some of the properties of Spanish verbal and nominal inflectional morphology. (The production and comprehension of verbal and nominal inflection in L2 Spanish are the focus of Chapters Four and Five, respectively.) In Section 2.2, I describe various aspects of

features: how generalizations can be captured using underspecification theory (2.2.1), and how morphological features interface with syntax (2.2.2, 2.2.3, 2.2.4); this section will establish a basis for the assumption of a realizationist, or "syntax-before-morphology" model of the morphology-syntax interface. I will assume (following Harley 1994, Harley and Ritter 2002, Carstairs-McCarthy 1998, among others) that underspecified features correspond to unmarked features; Section 2.3 establishes an independent basis, based on original data from Spanish, for the markedness relations I propose. Section 2.4 concludes this chapter with an overview of the theoretical framework that I adopt, Distributed Morphology (DM). DM allows for underspecification of features and provides a theoretical basis for the notion of *default*, something that will figure prominently in the chapters to follow.

2.1 Properties of Spanish inflection

I will begin this chapter with a discussion of the properties of verbs and nouns in Spanish, in particular, their associated morphological realizations of the abstract properties of tense, agreement, and finiteness. In Section 2.1.1 I discuss properties of the Spanish verb phrase (VP) with particular reference to the variables that I will examine in later chapters: tense, person, number, and finiteness. In Section 2.1.2 I discuss properties of the Spanish noun phrase (DP), with particular reference to gender and number agreement. Throughout the verbal and nominal morphology sections, I show cases in which the syntax appears to be sensitive to distinctions that are not reflected overtly in the morphology. These examples will serve as the basis for an argument in favor of underspecification; I present this argument in Section 2.1.3.

2.1.1 Spanish verbal morphology

Spanish verbs show many distinctions that are not overtly expressed in English. In this sense, Spanish is generally considered to be a language that has rich agreement, whereas English is a language that is morphologically impoverished. One task of the Spanish L2er is to acquire the various verb endings that spell out person, number, tense, and finiteness.

Verbs fall into three main conjugation classes that can be identified according to their infinitive markings: *-ar, -er*, and *-ir*. The vowel of each of these endings is often referred to as a *theme vowel* (*a, e, i*). The theme vowel links the root to the tense and agreement suffixes. Under some analyses this vowel is considered to be part of one of the attached suffixes (see for example Aguado-Orea 2004, cf. Montrul 2004). I will treat this vowel as part of a suffix, though it is not crucial to my analysis. The general format of verb morphology is illustrated in (1). The root is followed by one or more affixes that indicate tense/aspect/mood and person/number agreement.

1. *Suffixes in Spanish verbal morphology*

root	suffixes	word	
habl	ar	hablar	'speak (infinitive)'
habl	as	hablas	'you speak'
habl	aba, s	hablabas	'you spoke'

Not all of the categories that might be expressed overtly are expressed overtly. For example, *hablas* is a present-tense verb, but there is no present-tense marking *per se*. This contrasts with *hablabas*, which contains the past-tense marking *-aba*.

The person-number paradigms for person-number agreement are shown for the simple present indicative in (2) and the preterite past tense in (3). These paradigms correspond to a standard variety of pan-American Spanish[3]. By comparing (2) with (3), we can see that there is a wide variety of past-tense morphemes in Spanish, whereas in English, past is largely regular (as indicated by the suffix *–ed,* barring some exceptions like *put* and *dwelt*). In some cases, there is no overlap at all between past and present tense endings, as in 3rd person singular (*habla* 'he speaks' versus

[3] The main difference between Latin American and Peninsular Spanish agreement lies in whether there is an informal 2nd person plural (Peninsular Spanish *vosotros*) in use. None of the subjects in my study employed *vosotros*. Of course, this could be due to a lack of context for using the 2nd plural. On the issue of dialect, some regions of Latin America (e.g. Argentina) show different morphology for 2nd person singular agreement; however, none of the L2 subjects employed such forms. Finally, the majority of the teaching assistants at the university where this research was conducted come from Latin America; Ana Faure (p.c.) reports that Spanish teaching assistants from Argentina generally use the paradigm in (2). Thus I will assume that their target language contains the agreement patterns in (2).

habló 'he/she/it spoke'). In other cases, there is a total correspondence between past and present: *hablamos*, for example, can mean either 'we speak' or 'we spoke', as *–amos* corresponds to 1st plural for both tenses. The latter case is an example of *syncretism*—the appearance of a single form in more than one cell of a paradigm. (Syncretism, as we will see, is important in demonstrating the value of underspecification; I will return to this point in Section 1.3.).

2. *Spanish simple present indicative of ar/er/ir verbs*

	1st person	2nd person	3rd person
singular	o/o/o	as/es/es	a/e/e
plural	amos/emos/imos	an/en/en	an/en/en

3. *Spanish past preterite of ar/er/ir verbs*

	1st person	2nd person	3rd person
singular	é/í/í	aste/iste/iste	ó/ió/ió
plural	amos/imos/imos	aron/ieron/ieron	aron/ieron/ieron

Spanish has another past-tense, the past imperfect. The paradigms are broken down for this tense in (4). 1st and 3rd singular are syncretic, meaning that a verb like *hablaba* can mean either 'he/she/it spoke' or 'I spoke'. In addition, the *–er* and *–ir* class affixes are identical for this tense.

4. *Spanish past imperfect of ar/er/ir verbs*

	1st person	2nd person	3rd person
singular	aba/ía/ía	abas/ías/ías	aba/ía/ía
plural	ábamos/íamos/íamos	aban/ían/ían	aban/ían/ían

Comparing across the paradigms in (2-4), we can see that 2nd and 3rd plural are syncretic for simple present, past preterite, and past imperfect across all conjugations (though see footnote 2). Although Spanish shows rich inflection, syncretism is widespread in the verbal morphology system.

2.1.2 Spanish nominal morphology

Nouns in Spanish carry inherent gender properties—all nouns are either masculine or feminine. Other DP-internal elements such as determiners and adjectives agree in gender and number with the noun, as in (5)[4].

5.	a)	las	casas	nuevas
		the-FEM.PL	houses-FEM-PL	new-FEM-PL
	b)	el	libro	nuevo
		the-MASC.SG	book-MASC-SG	new-MASC-SG

I will begin this discussion by noting the properties of the nouns themselves. Plural number is indicated in (5a) by the *–s* morpheme on the noun, determiner, and adjective; singular is indicated in (5b) by a lack of plural marking. In general, plural marking is straightforward: either *–s* or *–Vs* (depending largely on the phonological properties of the stem) is affixed to a noun in the plural, although in some cases, no *–s* is added for plural marking (e.g. *las crises* 'the crises'; cf. *la crisis* 'the crisis'). For the most part, however, the alternation between plural and singular nouns is signaled by an affix for the plural, and a lack of affix for the singular.

The examples in (5) contain canonical Spanish nouns: *libro* and *casa* bear the canonical gender endings (*-o* for masculine and *–a* for feminine). There are many masculine nouns that do not end in *–o*, and likewise many feminine nouns that do not end in *–a*. Some of these are listed in (6).

[4] I will use the nouns *libro* and *casa* in a few gender examples from this point on in this chapter, without providing a gloss each time.

6.

masculine	*feminine*
lápiz 'pencil'	canción 'song'
coche 'car'	llave 'key'
mantel 'tablecloth'	piel 'skin'

Furthermore, there are many masculine nouns that end in *–a*, the canonical feminine ending. There is also a small number of feminine nouns that end in *–o*.

7.

masculine –a	*feminine –o*
problema 'problem'	mano 'hand'
mapa 'map'	modelo 'model'
comunista 'communist'	foto 'photo'

Although there is a tendency for *–o* to co-occur with masculine gender on one hand, and *–a* to co-occur with feminine gender on the other, there is no direct correspondence between gender and word ending, as many exceptions are found. These *–o* and *–a* endings cannot be said to "mark gender" in any reliable sense.[5] In Harris's (1991) terms, these endings are "word markers", not gender markers; see this work for an in-depth discussion of gender in the Spanish lexicon.

Determiners in Spanish agree in gender and number with the noun. The paradigm for definite (8) and indefinite (9) determiners is shown below.

8. *definite determiners*

	singular	*plural*
masculine	el	los
feminine	la	las

[5] Under some accounts in the L2 acquisition literature, these endings are claimed to introduce a gender feature to the word (see, for example, Franceschina 2001). Since *–o* and *–a* are unreliable at indicating gender, it cannot be the case that these affixes introduce gender to the word; if they did, we would expect no masculine *–a* words like *problema*, and no feminine *–o* words like *mano*. Under the view that I adopt later in this chapter, affixes do not introduce features, but rather realize features that already exist in syntax.

9. *indefinite determiners*

	singular	*plural*
masculine	un	unos
feminine	una	unas

The marking of the plural in determiners is signaled by *–s* or *–os* for the indefinite determiners and feminine definite determiners, but the contrast between *el* and *los* is suppletive—that is, the two forms are not plausibly phonologically related, as the contrast is not marked through a transparent process of affixation. Thus Spanish plural agreement shows some complexity when we consider the transparency of plural marking.

Spanish adjectives also agree in number and gender with the noun, as we saw in (5). Many adjectives such as *verde* 'green', however, do not show a gender distinction, as in (10); some do not show a number distinction either, such as *beige* 'beige' (11).

10.

	masculine	*feminine*
singular	el libro verde	la casa verde
plural	los libros verdes	las casas verdes

11.

	masculine	*feminine*
singular	el libro beige	la casa beige
plural	los libros beige	las casas beige

We have therefore seen another case of syncretism: *verde* is syncretic for gender, and *beige* is syncretic for both gender and number. Thus the generalization that emerges about gender and number agreement is that Spanish adjectives agree in gender and number, though we may not see this agreement in the case of some adjectives.

Throughout this section, we have seen cases of syncretism for several features, specifically: person agreement (e.g. 1st and 3rd person *hablaba* 'I/he/she/it spoke'), gender agreement (e.g. *verde* 'green-MASC/FEM'), number agreement (e.g. *beige*

'beige-MASC/FEM/SG/PL'), and tense (e.g. *hablamos* 'we speak/we spoke'). Thus we can conclude that morphology is not a reliable indicator of syntactic context: for example, we cannot always tell what the gender and number of a noun is by simply looking at the morphology of an agreeing adjective, nor can we reliably diagnose person agreement on a verb by simply looking at that verb's person agreement morphology. In the following section we will look at how this observation is captured in terms of morphological theory.

2.2 Features in morphology and syntax

In the previous section, we saw several instances of syncretism in which a distinction that could have been signaled overtly was not. In this section, I will begin by presenting one way in which syncretism has been modeled: via underspecification. I will also describe what an alternative account would look like: one that adopts full specification.

2.2.1 Underspecification

Underspecification is the notion that all information that is redundant can be excluded from a representation. It is also a tool used for dealing with the observation that the syntax may be sensitive to distinctions that are not systematically reflected in the overt morphology (Bobaljik 2002). I will be using underspecification to apply to a specific set of phenomena: those cases in which the syntax is specified for features for which there is no distinct overt morphological expression.[6] It is in the morphological component that underspecification applies, not in the syntax.[7]

I will start with an example involving gender agreement in nominative pronouns. Spanish distinguishes gender in nominative pronouns for 3rd person, but not 1st, as shown in (12).

[6] With this statement, I explicitly adopt the Separation Hypothesis (Beard 1995). This assumes that abstract morphosyntactic features are independent of their overt realizations.

[7] This differs from much of the L1 acquisition literature, in which underspecification is understood as a syntactic phenomenon (see, for example, Hyams 1996).

12. *Person/number in Spanish nominative pronouns*

	masculine	*feminine*
1	yo	yo
3	él	ella

There are two possible ways to represent gender in the pronoun *yo*. We could assume that, like the 3rd person pronouns *él* and *ella*, *yo* is specified for gender, but that we simply do not pronounce this difference. We then are forced to posit two separate lexical items that happen to be pronounced in the same way: one for *yo*-masculine and one for *yo*-feminine. In other words, we have a case of *accidental homophony* (in the same sense that *due* and *dew* are homophonous but have separate lexical entries, for example). This is the approach we would be forced to take under *full specification*. A problem for this approach is that it clearly misses a generalization: *yo*-masculine and *yo*-feminine seem to be the same lexical item. Alternatively, we could assume that there is a single lexical item that bears no specification for gender. This is what we would assume under a theory that adopts *underspecification*. One obvious advantage is that we do not need to posit more than one lexical entry for *yo*. I will assume, therefore, that underspecification is on the right track, in that it avoids having to posit multiple representations for the same lexical item, and therefore permits a more economical representation.

I will use the term underspecification to encompass two types of representation. The first is the type we have already seen for *yo*: gender is *unspecified*, which entails that neither masculine nor feminine is specified. The second type of representation involves the representation of contrastive feature values. As we have already seen, the pair of 3rd person nominative pronouns *él* and *ella* show a distinction for gender. Although both signal gender overtly, it is not logically required that both masculine and feminine be represented. Instead, we can eliminate one feature from the representation on the basis of redundancy. That is, the specification of one feature (e.g. feminine) can be paired with the underspecification of the opposed feature (e.g.

masculine) without the loss of any information, as the masculine interpretation can be deduced from the absence of a gender feature.[8] This is schematized in (13).

13.

pronoun	*gender features*
ella	[feminine]
él	[ø]

Under full specification, since all features are specified by definition, we cannot eliminate the masculine feature. Thus underspecification is valuable in eliminating redundancy from the grammar, and in capturing cases of syncretism, as we saw in the discussion following example (12).[9]

In the following section, I will discuss how the grammar works in conjunction with underspecification. For the sake of clarity, I will phrase the remainder of this discussion within one particular theory, Distributed Morphology (DM) (Halle & Marantz 1993). My adoption of this framework is motivated by its endorsement of underspecification (though in principle, other underspecification-based approaches may serve just as well).

2.2.2 Blocking and competition

If we accept that underspecification is a good way of dealing with syncretism, there is an additional assumption we need to make: there must be a mechanism to block the insertion of underspecified morphemes in contexts where they do not belong. This mechanism is known as the *Elsewhere Principle* (or Blocking Principle, or Panini Principle). One statement[10] of it comes from Halle (1997):

---

[8] For the moment, I will assume that masculine is underspecified relative to feminine. This decision will be motivated in Section 3.

[9] Of course, full specification is not fatal in itself: we could rely on an additional mechanism like a feature hierarchy that creates a hierarchical asymmetry between masculine and feminine, for example, to yield the same effects as underspecification.

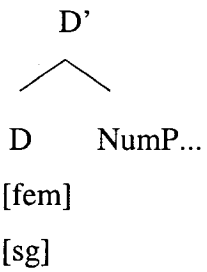
[10] Halle uses *vocabulary item* in the same sense that I am using the word *morpheme*, and *morpheme* in the sense of an abstract bundle of features on a terminal syntactic node, which I have been calling the syntactic context.

> "The phonological exponent of a vocabulary item [VI] is inserted into a morpheme... if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the VI contains features not present in the morpheme. Where several VIs meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen." (Halle 1997:128)

With this principle in place, it becomes impossible to use an underspecified morpheme in cases where there is a more specified one available that is consistent with the syntactic context. It is also impossible to insert morphemes with non-matching features—this would constitute an instance of *feature clash*. The syntactic context is a representation against which we evaluate the morphemes that could potentially be inserted; I will assume that the syntax—the context against which the morphemes are evaluated—is fully specified.

To illustrate how the Elsewhere Principle operates, consider the example of gender in (14), and the underspecification of masculine within a DP. In a feminine syntactic context, the most specified morpheme consistent with the context is a feminine determiner. Masculine does not clash, although it is blocked by virtue of the existence of a morpheme with a greater number of matching features. The remaining morpheme *el* can be inserted where no condition for insertion is met by any of the other competing morphemes. This operation is schematized below. In (14a), the syntactic context involves a DP with a fully-specified head D, containing the features [feminine] and [singular], ignoring any other features (like definiteness for example) for the sake of simplicity. In (14b), the competing morphemes (VIs) are shown along with their corresponding *rules of exponence*—that is, conditions for insertion. (I further assume that singular is underspecified, which I will motivate in Section 2.4). The winning VI is *la*, by virtue of the fact that it contains the greatest number of features consistent with the syntactic context, barring feature clash.

14. a) *syntactic context*



b)

competing morphemes	*outcome*
la ↔ [fem]	wins competition
el ↔ *elsewhere*	blocked

The Elsewhere Principle applies to regulate the competition for vocabulary insertion, preventing the overuse of underspecified morphology. We will see, in the following chapters, how this principle can be used to explain facts about L2 acquisition; specifically, L2ers appear to apply this principle in ways that are not entirely native-like.

2.2.3 The realization of features

Another crucial assumption underlying the operation of underspecification concerns the relationship between features and overt morphology. Specifically, we will see that the relationship between affixes and features is a complex one: the absence of an affix/affixes cannot be relied upon to signal the absence of a given feature or features when natural language data are considered. In this section, I will consider two views on the relationship between morphology and syntax: a *projectionist* approach that holds that morphology projects features into syntax, and a *realizationist* approach that holds that morphology is a mere realization of syntax and abstract features. The important point here is that underspecification is fully compatible with a realizationist approach, but is not straightforwardly compatible with a projectionist approach.

Under a projectionist approach, also termed "morphology-before-syntax" (e.g. White 2003), overt morphemes are the source of features. Returning to (5b), an adjective like *nuevas* is feminine and plural because of the addition of the *–a* and *–s* affixes. In a sentence like (15a), *alta* 'tall' is feminine because of the *–a* suffix. Agreement is calculated between the DP and the adjective and the best match wins; both bear a feminine feature. "Best match" rules out (15b) as a possible sentence; therefore, a projectionist approach that assumes underspecification is able to decide between (15a) and (15b), which contain the gender-specified pronoun *ella*.

15. a) [3][fem] [fem]
↑ ↑
ella es alt a

b) [3][fem]
↑ ↑
ella es alt o

In the next pair of examples, the projectionist approach breaks down. In (16), we have the pronoun *yo*, which we would assume to be underspecified for gender[11]. In the context of a feminine speaker, there is no way to decide between (16a) or (16b). Specifically, since no [feminine] feature is introduced by the syncretic pronoun, there is no way to reject (16b) when spoken by a feminine speaker, since neither is a match.

16. a) [1] [fem]
↑ ↑
yo soy alt a

[11] Of course, a projectionist approach could assume full specification here in order to get this derivation to proceed. However, recall that this assumption would mean that there are separate lexical entries for *yo*-masculine and *yo*-feminine. A projectionist approach does not "fail", but it is incompatible with underspecification. See note 9.

b) [1]

↑ ↑

yo soy alt o

Thus, a projectionist approach is not easily compatible with underspecification.[12] However, another approach, termed realizationist or syntax-before-morphology (e.g. White 2003), permits underspecification. Under this approach, affixes do not introduce features, but instead realize them. One particular realizationist approach, DM, holds that features are present in the fully-specified syntax, and the most highly-specified morphological form consistent with the syntactic context is inserted (according to the Elsewhere Principle).

To illustrate how a realizationist approach is compatible with underspecification, I will return to the pair of sentences in (16) that was problematic under a projectionist approach. *Yo* is the most highly-specified morpheme consistent with the syntactic context, which contains specification for the features [1][nom][fem] in (17-18); this is the case even though *yo* is not specified for gender. Similarly, the *–a* morpheme on the adjective *alta* is the most highly specified morpheme consistent with the feminine syntactic context, as shown in (17b).

17. a) [1][nom][fem] [sg][fem]

↓ ↓

yo soy alt-

b) *competing morphemes*

-a ↔ [fem] *wins*

-o ↔ elsewhere *blocked*

18. a) [1][nom][masc] [sg][masc]

↓ ↓

yo soy alt-

b) *competing morphemes*

-a ↔ [fem] *clash*

-o ↔ elsewhere *wins*

[12] More precisely, a projectionist approach requires additional mechanisms such as the paradigm-level to introduce features into syntax (see for example Wunderlich 1996). In principle, we could appeal to paradigms as extra levels of structure in order to get beyond the incompatibility of a projectionist approach with underspecification.

The projectionist approach fails because it relies upon morphology to introduce features that are necessary in the calculation of agreement; if these features cannot be introduced because of syncretism (which we assume to be represented via underspecification), the calculation of agreement fails. To sum, in this section we have seen that a projectionist approach, in contrast to a realizational approach, is incompatible with underspecification when the operation of agreement is considered.

2.2.4 Movement

In the previous section, I presented an argument against the claim that morphology introduces features to syntax; this must stand if we accept the claim that underspecification is a satisfactory way to capture syncretism. In this section, I will consider another claim that attributes a causal role to morphology: that the richness of morphological paradigms is a trigger for the syntactic operation of movement; as it turns out, this claim is similarly not supported. These arguments support a view of morphology as being a consequence of syntax, rather than the cause of syntax. In White's (2003) terms, these arguments support the "syntax-before-morphology" or realizationist approach to the syntax-morphology interface.

According to one influential view, the presence of overt morphology is claimed to trigger verb movement (Rohrbacher 1999, Vikner 1995). This claim is motivated by the relationship between the richness of morphological paradigms and verb raising. Data from two Germanic languages illustrate this point (from Rohrbacher 1999:1,ex. 1a,b).

19. Jón harmar að María keypti ekki bókina. (Icelandic)
    J. regrets that M. bought not book-the
    'John regrets that Mary didn't buy that book.'

20. Jag beklager att Eva inte köpte boken (Swedish)
    I regret that E. not bought book-the.
    'I regret that Eve didn't buy the book.'

In (19), the verb precedes negation; in (20), the verb follows negation. It is generally assumed in the Principles and Parameters framework (e.g. Chomsky 1981) that this word order is an effect of whether or not the finite verb has raised out of the VP to Infl, or some other functional projection above the VP (e.g. Pollock 1989). When the verb has raised to Infl, the result is the word order in (19). When the verb has not raised, the result is the word order in (20).

Icelandic and Swedish differ not only on word order, but on their use of overt morphological marking. Icelandic bears rich agreement morphology, whereas Swedish does not.[13] Below are some example paradigms in the indicative present.

21. *Person-number agreement in two Germanic languages*

a) Icelandic *kast-a* 'throw'

	singular	*plural*
1	köst-um	kasta
2	kast-ið	kasta-r
3	kasta	kasta-r

b) Swedish *smaka* 'taste'

	singular	*plural*
1	smaka-r	smaka-r
2	smaka-r	smaka-r
3	smaka-r	smaka-r

Icelandic shows rich agreement and verb raising to Infl; Swedish shows neither rich agreement nor verb raising. Typological evidence from the Germanic languages suggests that there is a correlation between the richness of agreement morphology and verb raising. This correlation has been drawn in two ways. Under one formulation, rich morphology is the cause of overt verb raising (Rohrbacher 1999, Vikner 1995). There is a bidirectional relationship between the richness of morphology and the presence of verb raising, such that a language has verb raising if and only if it has rich morphology. Spanish is like Icelandic in that it has rich morphology (see 2-4), and the option of verb raising over negation (in 22, the negative element *más*):

[13] Rich agreement is defined by Rohrbacher under the Paradigm-Verb Raising Correlate as follows: A language has V to I raising if and only if in at least one number of one tense of the regular verb paradigm(s), the person features [1st] and [2nd] are both distinctively marked (1999:116). Spanish marks 1st and 2nd distinctively and thus qualifies as a language with rich agreement under this definition.

22. Después de la crisis, Francisco no viajó más a Europa
After the crisis, F. NEG traveled-3SG NEG to Europe
(Montrul 2004:ex.4, p.92)

Upon closer examination, the bidirectional formulation of this relationship cannot be maintained (Sprouse 1998, Bobaljik 2002). Languages are found that have verb raising, but lack rich inflection. Below are examples from Afrikaans (23,24) from Conradie (2005), who argues that the language shows verb raising over negation, but does not show rich agreement. In fact, Afrikaans shows no person agreement whatsoever, as shown in (24), for the verb *werk* 'work'.

23. Sy ken nie daardie man nie. (Conradie 2005: ex. 46, p. 57)
she knows not that man final-NEG
'She does not know that man.'

24. *Afrikaans Present* (Conradie 2005: ex 52, p. 61)

ek ('I') werk	ons ('we') werk
jy ('you (sg)') werk	julle ('you (pl)') werk
hy / sy / dit ('he / she / it') werk	hulle ('they') werk

The facts in (23,24) suggest that the relationship is unidirectional: instead, we can adopt a second formulation of the agreement-movement relationship:

25. If a language has rich inflection, then it has verb raising to Infl. (Bobaljik 2002:132)

The existence of languages like Afrikaans thus provides typological evidence against a strong, bidirectional formulation of the relationship between morphology and syntactic operations. Thus we can see that the richness of morphological paradigms cannot be relied upon to diagnose movement, because an absence of morphology does not entail an absence of movement.

In Sections 2.2.3 and 2.2.4 we have seen that morphology and syntax are dissociated. In two senses, morphology does not drive syntax: morphology cannot be relied upon to introduce features, nor can the degree of richness of morphology be relied upon to diagnose movement. The implication, of course, is that if natural language grammars do not rely on overt morphology or the richness of inflection to introduce features or cause movement operations, we should not expect an L2 grammar to do so.

Given the generalizations outlined in this section, I adopt a realizationist or "syntax-before-morphology" approach, which I have shown to be compatible with underspecification. In the following section, I will adopt a particular view of underspecification that relies on markedness in the establishment of feature representations.

2.3 Markedness and underspecification

As a starting point in the inquiry into the representation of features, I will assume, following recent advances in morphological theory (e.g. Harley 1994, Bonet 1995, Harley & Ritter 2002, Carstairs-McCarthy 1998, Cowper 2004) that markedness drives the specification of features. Specifically, I will assume that underspecified features correspond to unmarked features. The focus of this section, therefore, is to establish exactly which features are marked and which are unmarked. I will consider some criteria for establishing markedness values in Section 3.1. In Section 3.2, I will apply these criteria to data from Spanish. In Section 3.3, I summarize the feature representations I assume, and define the notion of *default*, something that will be crucial in supporting the use of underspecification.

2.3.1 Markedness: Background

Markedness has been employed in various senses. For example, unmarked forms are thought to be those that are more basic or general, use less structure, are acquired first, and are typologically more frequent, whereas marked forms are more complex, use more structure, are acquired late, and are typologically more rare (Battistella 1990). There is considerable difficulty in defining this term. In this section, I will

focus on markedness as it is defined in terms of meaning and distribution. Thus the focus will be on the meanings of morphology, rather than on morpho-phonology.

Markedness theory deals with oppositions in language. From early studies in structuralist linguistics (e.g. Jakobson 1957/1984), we know that oppositions in language are very often not best characterized as mere opposites. Instead, one term within an opposed pair is more general than the other. The more general term is unmarked, whereas the more complex term is marked. Thus unmarked forms have a dual function: they can mean the opposite of a marked term, but they can also mean the absence of signalization of the marked term.

To illustrate, I cite an example from tense that comes from Battistella (1990:3-4). This example shows that past and present tenses do not behave as mere opposites. Instead, past tense unambiguously signals past time, whereas the present tense is not necessarily specified for time. In (26), present tense demonstrates a range of temporal meanings: it is used to indicate a habitual reading that is independent of time (a), to signal the future (b), and to signal the past (c).

26. a) I wear sneakers. (3:ex. 3b)
    b) I arrive home Sunday (3:ex. 3c)
    c) So then I say to him, "Shut up!" (4:ex. 3d)

The same pattern extends to Spanish preterite past as well. The sentences in (27) below are Spanish translations of (26); native speaker interpretations correspond to the time readings in the (26) examples.

27. a) Llevo zapatillas.
       Wear-1SG sneakers
       'I wear sneakers'
    b) Llego a casa el domingo.
       Arrive-1SG-PRES at home the Sunday
       'I arrive home on Sunday'

c) Y luego le digo, "¡Cállate!"
And later to-him say-1SG "be-quiet-2SG-REFL"
'And later I say to him, "Shut up!"'

Compare (27) with (28) below. These examples are unambiguously read as past tense.

28. a) Llevé zapatillas.
Wear-1SG-PAST sneakers
'I wore sneakers'
b) Llegué a casa el domingo.
Arrive-1SG-PAST at home the Sunday
'I arrived home on Sunday'
c) Y luego le dije, "¡Cállate!"
And later to-him say-PAST.1SG "be-quiet-2SG-REFL"
'And later I said to him, "Shut up!"'

These examples from English and Spanish illustrate that present and past are not mere opposites. Instead, the present tense is the general tense in that it is better understood as the non-specification of tense. These examples suggest that present is unmarked relative to the past tense.

2.3.2 Representing markedness: Bundles and feature geometries

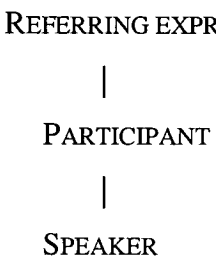
Markedness relations are frequently encoded in morphological and phonological theory through underspecification, with marked forms bearing additional features or additional structure in order to capture asymmetrical relations between features (e.g. Bonet 1995, Harley 1994, Harley & Ritter 2002, Cowper 2004, among others). The specific proposals vary: for some authors, features are represented as bundles (e.g. Noyer 1997); for others a feature-geometric structure is adopted (e.g. Bonet 1995, Harley 1994, Harley & Ritter 2002). The "bundle" approach is the type I have shown in (13): the bundle corresponding to a feminine pronoun (or determiner or clitic, etc.)

contains a feature [feminine], whereas the bundle corresponding to a masculine determiner contains no specification for gender.

A feature-geometric approach holds that features are hierarchically ordered rather than simply listed or bundled. Under this type of representation, features enter into structural dependencies such that the distribution of certain features is linked to that of others. Natural classes such as person and number are characterized as organizing nodes that dominate daughter nodes—the features that fill in the members of these natural classes—so that the presence of a node such as 'plural' entails the presence of the dominant, natural-class node of 'number'. Harley and Ritter (2002) note that although these dependencies must be a universal property of language, morphological theory has often failed to attempt to characterize them, as features are often represented as unstructured, unorganized bundles. Feature-geometric representations, on the other hand, capture these natural classes and structural dependencies in a straightforward way.

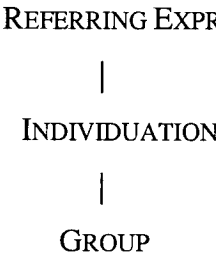
To illustrate how a feature-geometric approach works, I will cite the example of person and number features. Features are encoded through the presence or absence of a PARTICIPANT node (to indicate $1^{st}/2^{nd}$ person versus 3^{rd} person) and its dependant, SPEAKER. 1^{st} person is encoded through the presence of PARTICIPANT and SPEAKER. 2^{nd} person is encoded through the presence of PARTICIPANT and absence of SPEAKER. The presence of SPEAKER entails the presence of PARTICIPANT, as SPEAKER is the dependant of PARTICIPANT. The absence of PARTICIPANT (and, correspondingly, of SPEAKER) encodes 3^{rd} person. Markedness is encoded in that 1^{st} person bears the most structure and is the most marked, and 3^{rd} person bears the least structure and is unmarked. This is represented in a partial geometry in Figure 1, extracted from Harley and Ritter (2002).

Figure 1. A feature-geometric approach to person



Similarly, Harley and Ritter represent number through the feature GROUP. In a language such as Spanish that encodes only a contrast between singular and plural (as opposed to encoding singular/dual/plural, for example), only one feature is needed (INDIVIDUATION serves to mark the mass/count distinction, which I will put aside). Singular is represented via the presence of a bare INDIVIDUATION node, and plural via the presence of a GROUP node that is dominated by INDIVIDUATION.

Figure 2. A feature-geometric approach to number



Gender can be represented the same way, as noted by Harley (1994), who represents FEMININE as a dependant of GENDER; MASCULINE is represented via a bare gender node. Cowper (2004) similarly outlines a feature-geometric account of tense and aspect in Spanish, where the presence of PRECEDENCE encodes past, and its absence encodes nonpast. The issue of how features are best represented, via geometry or via bundle, will not be directly addressed until Chapter 6. Instead, I will focus on what these two representation schemes have in common: the assumption of a link between markedness and the quantity of representational structure, be it feature or node. Under

both representation schemes, an asymmetrical relationship between marked and unmarked features is crucially encoded.

Thus, following these authors, I make the assumption that unmarked equals underspecified. In order to motivate the feature inventories I will be adopting, I will establish an independent basis for arguing that certain features are unmarked (and, by assumption, underspecified) as opposed to other features that are marked. There are several criteria outlined in the literature for establishing markedness values between terms in opposition. In the sections below, I consider some of the criteria that have been posited in the literature. I will focus on semantic and distributional criteria, keeping in mind that no criterion is 100 percent failsafe; in all cases more than one criterion will be applied. The criteria given below follow the terminology and logic of Battistella (1990), with original examples from Spanish to determine markedness values at a language-specific level.

2.3.3 Criteria for establishing markedness relations

Indeterminateness is a semantic criterion that holds that a marked element has a specific meaning, while the unmarked element is indeterminate. The unmarked term has a general interpretation, and so it may substitute for the marked term in some contexts (Battistella 1990:27). The use of a present-tense verb to mean past-tense as in (27c), and the corresponding impossibility of a present-tense meaning of past-tense verbs (28) suggests that present tense is unmarked relative to past.

Neutralization is a distributional criterion; it occurs when a marked term is excluded from some context in which an unmarked term can occur (Battistella 1990). For example, in English, there are many pairs of words of which one member is specified for feminine (*actress*, *lioness*), while the other member is not necessarily specified for masculine (*actor*, *lion*). If we look to a plural context, the category *lions* can occur in the context of describing a group of masculine and feminine felines; the feminine *lionesses* cannot occur in such a context. The broader distribution of *lions* over

lionesses suggests that *lion* is unmarked. As we will see, the same pattern of results is found for gender in Spanish as well.

Syntactic distribution is a cover term that can refer to a few different criteria; here, I take it to mean the "occurrence in a wider range of contexts in a language" (Battistella 1990:38). The indeterminacy of the unmarked term results in its use in a wider range of contexts; in some contexts, there is no basis for determining a meaning difference between two values, and so this criterion is not the same as indeterminateness. An example is the use of a masculine determiner to introduce a borrowed word in Spanish (e.g. *el fax*).[14]

Syncretization addresses the use of formal distinctions in morphology, and is defined as "the elaboration of a category by a greater or fewer number of subdistinctions" (Battistella 1990:27). This criterion follows an observation of Greenberg (1966) that the unmarked term within an opposition shows more distinctions (e.g. gender, number) than do marked terms. For example, 3rd person pronouns are much likelier than either 1st or 2nd person to show distinctions based on gender, location, and class (Forcheimer 1953), suggesting that 3rd person is the unmarked person. This point is elaborated in (33) below.

Formal marking is "the relation between two opposed units of linguistic expression such that one is characterized by an augmentation, compounding, or complexity of form that the other lacks" (Battistella 1990:34). One example of formal marking is affixation: *dog* is unmarked relative to *dogs* due to *–s* affixation.

This correlation between formal marking and markedness is far from absolute, yet Greenberg (1966), among others, uses formal marking to establish markedness relations. In the theoretical literature, Carstairs-McCarthy (1998) has taken the correlation between formal marking and markedness to its extreme, arguing that an inflectional marking can never have the meaning of an unmarked value like singular

[14] There are actually a number of factors that play a role in determining how gender is assigned to borrowed words by bilingual speakers, including the phonological shape of the word and the biological gender of the referent (see, for example, Barkin 1980).

or present tense (and hence an affix cannot be specified for an unmarked feature like singular or present). However, many instances are found in which the unmarked category—established on semantic or distributional grounds—shows overt marking that is equal to or more complex than the marked category. To take an example from English, 3rd person and singular have been argued on semantic grounds to be least marked, which is to be justified in Section 2.4, yet English present tense shows marking only in 3rd person singular (*he sees*). Thus Battistella (1990) concludes that the correlation between formal marking and semantic markedness is a tendency at best, and rejects this as a criterion for establishing markedness values. I will follow him in avoiding formal marking as a criterion.[15]

2.4 Markedness relations

In this section, I apply the markedness criteria proposed in Section 2.3.3 to the variables I investigate in L2 Spanish: gender, number, person, tense, and finiteness. The goal of this section is to establish through markedness relations which feature within an opposition is unmarked and therefore, by assumption, underspecified.

2.4.1 Gender

Neutralization suggests that masculine gender is unmarked in Spanish. The plural of *hermano* 'brother' is *hermanos* 'brothers/siblings', which can include male and female siblings; the plural of *hermana* 'sister', *hermanas*, can only include female siblings.

Syntactic distribution indicates that masculine is unmarked. Harris (1991) has argued, on the basis of facts about the Spanish lexicon, that masculine is underspecified in Spanish. Masculine agreement surfaces in a wider range of syntactic contexts, including those in which no trigger for agreement is present. One example that supports his claim is the use of the preposition *para* 'for':

[15] A central point of this dissertation is that features are relevant—the "meaning" component of the sound-meaning connection in language. I take the position that it is the features that matter in determining the outcomes of L2 variability, not necessarily the morpho-phonological forms. Thus it seems invalid to try to adopt a diagnostic that appeals to morpho-phonological forms in establishing markedness values.

29. Tienes demasiados "paras" en este párrafo (Harris 1991:43)
Have-2SG too.many-MASC paras in this paragraph.

Since *para* is a preposition, it has no gender and cannot transfer gender to the quantifier *demasiados*. There is no alternative source of masculine gender in this sentence, so this presents a clear argument in favor of masculine gender as the default, at least in the case of Spanish.

2.4.2 Number

The meaning of singular number is indeterminate. If we consider the following pair, it becomes clear that singular is unmarked in Spanish relative to the plural.

30. a) El murciélago es nocturno.
the.SG bat is nocturnal
✓ intended meaning singular: one bat is nocturnal
✓ intended meaning plural: bats in general are nocturnal

b) Los murciélagos son nocturnos.
the.PL bats are nocturnal-PL
* intended meaning singular: one bat is nocturnal
✓ intended meaning plural: bats in general are nocturnal

In (30a), the morphologically singular *el murciélago* is able to take on a plural meaning. The plural, however, has an unambiguously plural meaning.

Syntactic distribution indicates that singular is unmarked. The question word *quién* 'who' is used when it is not known whether the answer is singular or plural. *Quienes* 'who-PL', however, is restricted to those contexts in which the speaker presupposes that the answer is plural. This contrast is shown in (31).

31. a) Quién comió las galletas? (expected answer: singular or plural)
who ate-3SG the cookies

b) Quienes comieron las galletas? (expected answer: plural)
who- PL ate-3PL the cookies

These examples also suggest that the meaning of *quién* is indeterminate.

2.4.3 Person

Indeterminateness indicates that third person is unmarked. Indefinite noun phrases in the third person, such as *uno,* can be used to refer to any person, including the speaker and hearer.

32. Uno debe pedir permiso.
one should-3SG ask-INF permission

Syncretization also indicates that 3rd person is unmarked. Greenberg (1966) notes that unmarked values tolerate more distinctions than marked ones. Typologically, 3rd person is more likely than 1st or 2nd person to show gender/number distinctions, suggesting that 3rd person is unmarked. This is true for Spanish, as shown in (33).

33. *Spanish singular nominative pronouns*

	masculine	feminine
1	yo	yo
2	tú	tú
3	él	ella

Spanish does not show gender distinctions in first or second person singular nominative pronouns, but it does show gender distinctions in 3rd person. Thus the language-specific evidence from Spanish supports the generalization that 3rd person tolerates more distinctions than other persons, suggesting that 3rd person is the unmarked person.

2.4.4 Tense

The particular contrast I will be dealing with is past versus nonpast/present tense. The examples in (27,28) show that present tense is indeterminate, and therefore unmarked.

Further evidence comes from syncretization: Spanish shows more person distinctions in present tense than past imperfect. Present distinguishes 1st person from 3rd, whereas the past imperfect makes no such distinction.

34. *Present and past imperfect singular of* hablar

1	hablo	hablaba
2	hablas	hablabas
3	habla	hablaba

The past preterite, however, makes an equal number of person distinctions as the simple present, as shown in (2-3). This would suggest that the past preterite and simple present are less marked than the past imperfect, and that the past preterite and simple present are equally marked. The criterion of indeterminateness suggests otherwise, however, indicating that past is marked relative to present. The markedness between the two past aspects, preterite and imperfect, will not be dealt with here, but might be an issue for future research.

Thus indeterminateness and syncretization suggest that present tense is unmarked relative to past.

2.4.5 Finiteness

On the basis of indeterminateness, nonfinite verbs are unmarked. Finite verbs are distinguished from nonfinite verbs on the basis of their specification for tense and/or mood. Infinitives, however, are unspecified for tense, and in subordinate clauses, they are dependent upon the matrix clause for temporal interpretation (Battistella 1990:106). In (35), the temporal interpretation of the infinitival clause *cerrar la puerta* 'close the door' differs according to the tense of the matrix clause.

35. a) Ana se acordó de [cerrar la puerta].
Ana remembered [to close-INF the door].
b) En cinco minutos, Ana va a [cerrar la puerta].
In five minutes, Ana will [close-INF the door].

2.5 Conclusion: Summary of features and markedness relations

In this chapter, I argued that underspecification is a useful tool in capturing the fact that not all syntactic distinctions are morphologically encoded. Following insights in theoretical morphology, I assume that underspecified features are those features that are unmarked. Using markedness criteria, I established markedness values for the variables under investigation. The markedness criteria have revealed the following relations for Spanish morphology:

36.

	Unmarked	*Marked*
Gender	masculine	feminine
Number	singular	plural
Person	third	non-third (1st or 2nd)
Tense	present (nonpast)	past
Finiteness	nonfinite	finite

I will therefore assume that the features in the unmarked column are not available in lexical (vocabulary) items.

In this chapter, I have used markedness criteria to establish an independent basis for feature inventories in Spanish. The following chapter shows how these feature inventories, when employed within a realizational framework like DM, allow us to set up a theoretically-based definition for the term *default*. *Default* is used in L2 acquisition literature as a descriptive term; when errors are attested in tense or agreement morphology, the morpheme the learner employs is described as a default.[16] I will begin the following chapter with a discussion of the literature on errors in

[16] The use of the term *default* is highly frequent in the L2 morphology literature. See, for example, White et al (2004), Bruhn de Garavito (2003a,b), and many others.

morphology in L2 acquisition, and we will see that although systematic substitution errors are found in previous work on L2 morphology, there is currently no theoretical motivation that *explains* why certain defaults might surface as opposed to others. The following chapter lays out my proposal, the Morphological Underspecification Hypothesis, which attempts to take a step toward explaining these facts in L2 acquisition.

Chapter 3
Morphological Variability in L2 Acquisition: Generalizations and Explanations

It is a well-attested fact that second language learners (L2ers) do not consistently produce the overt morphemes associated with tense and agreement. There is no consensus on what morphological variability means, or why it exists. Variability is attested in a range of L2s and across various levels. Research on variability encompasses a variety of L1s, L2s, and levels of proficiency, yet in spite of these differences, several generalizations emerge. Section 3.1 of this chapter will outline some of the basic generalizations regarding variability that have been reported in the L2 literature, and discuss how two different approaches have attempted to account for this variability. Section 3.2 outlines the approach advocated in this dissertation. Section 3.3 compares this approach to alternative proposals on morphological variability.

3.1 An overview of the phenomenon

Early studies on L2 morphology focused on the establishment of morpheme acquisition order, following Brown's (1973) work on L1 English. The goal of these studies was to show that the acquisition of L2 English had certain immutable characteristics that were independent of the L1 of the learner (e.g. Dulay & Burt 1974). More recently, the focus has shifted from the order of acquisition to the cause of morphological variability: when learners fail to (correctly) use inflectional morphology, including free morphemes like determiners and auxiliaries, what does this imply about the interlanguage grammar?

One view holds that morphological variability results from an underlying representational deficit in the syntax. Depending on the specific theory, this deficit may be either temporary or permanent. Vainikka and Young-Scholten (1994), for example, propose that the L2 initial state has only lexical categories, and lacks

functional categories in general. As functional categories and projections gradually emerge, inflection does as well. Vainikka and Young-Scholten take the inconsistent use of inflection to mean that functional categories have not yet emerged, or that the learner is in a period of transition from a stage that lacks functional categories to a stage that has functional categories. Thus variability is a developmental phenomenon that will eventually be overcome, as it is in L1 acquisition. Under another view, morphological variability is indicative of a permanent impairment in the grammar. Clahsen (1988), for example, argues that L2ers are unable to acquire the triggering relationship between morphological paradigms and verb movement, which he assumes child learners are able to acquire. (This view explicitly assumes a projectionist approach to the relationship between verb movement and overt morphology, something that I presented in Chapter Two as problematic.) For Clahsen, the inability to acquire this relationship is a consequence of a lack of availability of UG in L2 acquisition. Meisel (1991) argues for a similar contingency between the acquisition of overt morphology and verb raising; for both Clahsen and Meisel, L2 learners cannot acquire the necessary contingency between overt morphology and verb movement, and therefore suffer from a permanent impairment.

Hawkins and Chan (1997) argue that variability is reflective of an underlying permanent inability to represent uninterpretable L2 features[17] that are not instantiated in the L1; this is the essence of the Failed Functional Features Hypothesis (FFFH). To take an example from L2 Spanish (from Franceschina 2001, Hawkins & Franceschina 2004), L1 English speakers do not have gender as a functional feature in their L1, and thus cannot represent it in their L2. According to the FFFH, L2 Spanish speakers should perform poorly on gender agreement when they come from an L1 such as English; L2ers from another L1 that has gender would be expected to perform better.[18] Morphological variability—in particular, the lack of consistent use of overt

[17] For Hawkins, the problem lies in uninterpretable features, that is, those features with a purely syntactic function, such as agreement features on a verb. Uninterpretable features are those that must be checked during the course of a syntactic derivation. These contrast with interpretable features, which carry a semantic interpretation (such as agreement features on a pronoun), and which Hawkins holds to be unproblematic.

[18] As it turns out, they do not necessarily perform better. White et al (2004) show no difference between L1 French and L1 English groups in the acquisition of Spanish gender. I discuss this study in greater detail in Chapter 5, which deals with the acquisition of gender and number.

markers—is interpreted as evidence for the lack of underlying features, under the FFFH. For example, Franceschina (2000) takes the persistent variability in gender agreement by a near-native L2 Spanish speaker to mean that gender is not acquirable. Hawkins and colleagues, in other words, assume that variability in overt morphology tells us something about underlying syntactic competence.

Recently, however, the equation of missing or incorrect morphology with syntactic deficits has been called into question. Crucially, Lardiere (1998a,b) shows that the absence of (correct) overt morphology does not entail corresponding syntactic deficits in the L2 grammar. In a case study of Patty, a near-native speaker of L2 English (L1 Chinese), Lardiere (1998b) showed that the absence of overt tense morphology does not entail the absence of TP, if this is assumed to be the locus of nominative case assignment. Patty exhibited perfect nominative case assignment, but impaired use of tense morphology. The rate of use of past-tense morphology was approximately 34 percent in obligatory contexts. An example of nominative case (in the form of the pronoun *they*) paired with missing tense (in the form of a missing copula) is shown in (1).

1. Yesterday they open until five. (Lardiere 1998a: 16)

Patty also shows robust evidence for a CP projection, as shown in (2). If we accept that the presence of CP entails the presence of all the projections below it (following Grimshaw 1994), Patty must have the TP projection as well. The embedded CP is shown in brackets in (2).

2. a) but I know [$_{CP}$ that I have doubt] (Lardiere 1998a: 19, ex. 2e)
   b) I think [$_{CP}$ that we are so lucky] (Lardiere 1998a: 19, ex. 2h)

Lardiere (2005) captures the dissociation between syntax and morphology through the notion of *morphological competence*. L2ers are able to acquire syntactic structure, contra the FFFH, but run into problems in recognizing which features are activated

and bundled together in the target language, and how they are subsequently spelled out.

Prévost and White (2000b) similarly find that missing overt inflection does not entail an absence of verb raising or the functional projections (IP/AgrP) and feature strength (which presumably drives verb raising) associated with these projections. L2 learners of French and German have at least partial knowledge of verb placement with respect to negation (see Section 3.1.2). Finite verbs are found in raised positions (above negation), as shown in (3a): finite raised *peut* and nonfinite unraised *dormir*, and (3b): finite raised *mange*. In addition, nonfinite verbs were sometimes raised to finite positions, suggesting a dissociation between the abstract feature strength associated with verb movement on one hand, and the overt manifestation of tense and agreement on the other. The acquisition of the feature strength that triggers verb movement runs counter to the predictions of the FFFH.[19]

3. a) mais on peut pas dormir (PW 2000b:117, ex.14a)
but one can-1/2/3S not sleep-INF

b) i mange pas (PW 2000b:117, ex.14b)
he eat.1/2/3S not

Similar evidence of a dissociation between case and overt morphology is reported by Haznedar and Schwartz (1997), who find perfect case assignment alongside variable tense and agreement morphology in a Turkish child learning L2 English.

Together, these results suggest that the absence of agreement and tense morphology in L2 does not entail syntactic impairment. Furthermore, as I showed in Chapter Two, typological evidence argues against the theoretical claim that the richness of overt morphology drives verb raising: languages are found in which verb raising is allowed, but morphology is impoverished, thereby demonstrating that the relationship between overt morphology and verb raising cannot be bidirectional. This

[19] In light of evidence (from Parodi et al 1997) that L2ers produce correct word order in the DP very successfully even when it differs from L1 word order, Hawkins (2001:254) concedes that L2ers may be able to acquire new feature strengths that account for the difference in word order. However, as White (2003:127) points out, it is not clear that the FFFH should be able to distinguish between features and feature strength.

generalization is captured for L2 acquisition under the Missing Surface Inflection Hypothesis (MSIH; Prévost & White 2000b; see also Haznedar & Schwartz 1997), which holds that syntactic structure may be unimpaired despite the presence of morphological problems. Thus features, feature strength, and syntactic projections are present, but their overt morphological realizations are sometimes omitted in production. Under the MSIH, communication pressures are cited as an explanation for variability in production. While Prévost and White (2000b) and Lardiere (1998a,b, 2000, 2005) agree that syntactic representations are intact, the problems for the learner are perhaps seen as more profound than mere production problems under Lardiere's approach (see Section 3.3 below for further discussion).

To summarize, one generalization we can draw from the L2 literature is that there is frequently a dissociation between the (often incorrect) use of morphology, on one hand, and the knowledge of syntax, on the other. The variable use of inflection has, to this point, been almost exclusively examined from the point of view of syntactic effects that may or may not be related to morphological variability. In the remainder of this chapter, I will focus on generalizations that pertain to morphology in itself, rather than on generalizations that relate to facts about morphology's relationship to syntax. Section 3.1 considers one possible generalization: that morphological variability relates to the presence versus absence of morphology. Under this view, L2ers go through a period during which missing morphology alternates with present morphology (and may in fact never move beyond this period, and instead remain fossilized). Section 3.2 considers another generalization: that morphological variability involves alternations between correct and incorrect morphology. The second generalization, as it turns out, is supported over the first. Nevertheless, both types of cases— those that involve missing inflection and those that involve incorrect inflection— share one property: the use of systematic defaults. In both Sections 3.1 and 3.2 I note available data that supports this conclusion. In Section 3.3 I discuss task effects on morphological variability, drawing comparisons between studies of production and studies of comprehension, where the latter are available. A comparison of variability across tasks will tell us whether the phenomenon is strictly production-based and therefore potentially non-representative of the underlying

competence, or if there is a comprehension deficit that is potentially suggestive of a deeper source.

3.1.1 The presence versus absence of morphology

One generalization that emerges from the literature on L1 morphological variability is that when morphology is supplied, it is accurate (e.g. Poeppel & Wexler 1993 for German; though see Aguado-Orea 2004 for a different conclusion based on child Spanish). This generalization has held for much L2 research as well, particularly when it involves L2 English (e.g. Lardiere 1998a,b; White in press). In this section, I discuss some of the L2 data that has been offered in support of this generalization. We will see, however, that this generalization does not seem to hold for L2 acquisition in general; the presence of morphology does not entail the correctness of morphology.

In the acquisition of L2 English morphology, the "if present, then accurate" generalization clearly holds for two pieces of bound morphology: 3rd person singular *–s* and past-tense *–ed.* These morphemes are essentially never used in contexts in which they are not appropriate. 3rd person *–s* is frequently omitted from 3rd person contexts, as is past-tense inflection in past-tense contexts; White's (in press) data on suppliance versus "oversuppliance" of morphology (that is, the use of *–s* or *–ed* or corresponding irregular morphology in contexts where a bare verb is appropriate) are presented in Table 1. The distribution of these affixes suggests that the presence of overt affixes, at least for these cases, entails their correctness.

Table 1. Suppliance of 3rd singular–*s* and past-tense versus oversuppliance in L2 English (adapted from White, in press)

	3rd singular –*s*		**Past**	
	Suppliance	**Oversuppliance**	**Suppliance**	**Oversuppliance**
L1 French	60%	0%	50% reg/ 83.5% irreg	1.55% (reg and irreg)
L1 Mandarin	30.9%	0%	48% reg/ 83.33% irreg	5.25% (reg and irreg)

The above data suggest that, in the case of English, the default verb form is clearly the base/uninflected verb, as in (4), where past *–ed* is missing. (4) comes from a speaker of L1 Mandarin.

4. I never saw them before; they open my brain (White in press: ex. 16a)

Further examples of the base/uninflected verb as a default come from Patty (Lardiere 1998b). Patty produces *–s* in obligatory contexts surfaces at a rate of less than 5 percent in main verbs, considerably less than the L1 French/Mandarin speakers shown in Table 1, despite the fact that her proficiency level is very high. An example of missing –s is shown in (5). As discussed above, Patty frequently omits *–ed* in obligatory contexts; an example of missing *–ed* is shown in (6).

5. because he understand better now (Lardiere 1998b:368)
6. he call me last night (Lardiere 1998a:18)

Morphological variability is also attested in nominal domains. Errors involving missing determiners are shown in (7-9). In (7), a speaker of L1 Mandarin omits the indefinite determiner *a*. In (8), plural *-s* is omitted from the quantifier *demasiado* 'too many'; intermediate L2 Spanish learners omitted the *–s* on quantifiers in a written

production task at a rate of about 5 percent, but presumably do not oversupply *–s* (Español-Echevarría & Prévost 2004). Thus, for the examples we have seen thus far, the default is a zero morpheme.

7. and she made phone call to someone (White in press: ex. 19a)
8. demasiado libros (Español-Echevarría & Prévost 2004: 163)
   too.many.SG books-PL

A crucial example against the generalization that morphological errors typically involve the omission of inflection comes from L2 Dutch gender; although we see an alternation between a zero morpheme and an overt marking, the default is *not* a zero morpheme. Blom and Polisenska (2005) and Sabourin (2003) both report on the overuse of common gender in neuter contexts, suggesting that common gender acts as a default in determiners and adjectives. The contrast between common and neuter adjectives involves the presence versus absence of an overt morpheme, as shown in (9-10): *–e* corresponds to common gender, and zero corresponds to neuter gender. Sabourin's data come from a grammaticality judgment task in which subjects were given sentences like the one in (11), with either the correct or the incorrect adjective.

9. een klein-e tafel (Sabourin 2003:49)
   a small-COM table.COM
10. een klein kind (Sabourin 2003:50)
    a small.NEUT child.NEUT
11. Hij loopt op een gekke/*gek manier. (Sabourin 2003:50)
    he walks in a funny-COM/*funny-NEUT way.COM
    'He walks in a funny way.'

Sabourin's subjects were more likely to accept common *–e* in contexts in which it did not belong (e.g. **een kleine kind* 'a small-COM child.NEUT') than to accept a missing *-e* in contexts in which it should have occurred (e.g. **een klein tafel* 'a small.NEUT table.COM'). Blom and Polisenka's data come from an elicited production task by

Moroccan children of various ages acquiring Dutch as an L2, and similarly show much higher rates of overuse of *–e* than of missing *–e* (78-85 percent versus 7-8 percent, respectively).

Default common *–e* is therefore a crucial example, as it illustrates the inadequacy of an account based solely on the absence of morphology. Put simply, defaults do not equal zeros. Furthermore, these studies on L2 Dutch also find that subjects use common gender *de* as a default determiner in place of the neuter *het*, suggesting that the source of the problem may lie in the features these morphemes realize, rather than in the use versus non-use of morphemes.

In sum, the "if present, then accurate" generalization cannot be maintained given the Dutch example we have just seen. Although this generalization seemed to capture the facts regarding L2 English quite accurately, it is an oversimplification to say that the problem for acquirers of all L2s lies in *supplying* morphology. In Section 3.1.2, I present more examples that contradict this generalization, and attempt to draw connections among these examples. I will conclude by noting that affixation per se cannot be the cause of the problem. Instead, I will argue that the problem involves the feature content behind the overt realizations (or lack thereof) of morphology.

3.1.2 The use of incorrect morphology

In L2 German and French, case studies have found that the nonfinite verb,[20] as in (12-13), surfaces as a default:

12. monsieur il arriver (PW 2000b: 124)
    mister he arrive-INF
13. möchten ma du ein Kaffee? (PW 2000a)
    want-INF then you a coffee
    'Would you like a coffee?'

[20] Though see Müller (1998) for a counterexample—in a case study, an L2 German speaker used infinitival *–en* and (1st) person *-e* as a default. She notes that in some dialects of German, *-e* is the infinitival ending. Meisel (1991) also reports on the substitution of person forms in German L2, though the speech of only a few speakers make up the data set, so it is not clear how robust of a phenomenon errors in person might be.

The errors in (12-13) are of a different sort from the L2 English and Spanish errors we have seen in Section 3.1.1. The alternations in (4-8) involve a default that is missing an affix, whereas the alternations in (12-13) do not. The default in (12) contains more morphology than the target, and different morphology in (13). The defaults and targets for these two examples are shown in (14).

14.	*default employed*	*target*	
	arriver (arriv[e])	arrive (arriv[ø])	L2 French: PW 2000b
	möchten	mochst	L2 German: PW 2000b

Errors involving the overuse of infinitives in finite contexts are much more frequent than errors involving the overuse of finite verbs in nonfinite contexts in the PW data set, consisting of data from 4 L2 subjects, shown in Table 2 below. Subjects overextend the marking that corresponds to a nonfinite verb to contexts in which it does not belong; they generally do not overextend finite markings, however. The systematic use of overt nonfinite morphology in place of targetlike finite morphology suggests that nonfinite verbs may have some kind of default status in the interlanguage grammar. Citing inappropriate use of overt morphology in the shape of infinitive markers, White (in press) points out that "when morphology is present, it is not necessarily appropriate" (p. 2).

Table 2. Overuse of nonfinite morphology vs. overuse of finite morphology (Data from PW 2000b: 119)

	Obligatory finite contexts		Obligatory nonfinite contexts	
L2 French Abdelmalek	767	243(24.1%)	278	17(5.8%)
L2 French Zahra	755	224(22.9%)	156	2(1.3%)
L2 German Ana	389	45(10.4%)	76	7(8.4%)
L2 German Zita	434	85(16.4%)	98	6(5.8%)

Similar evidence from Spanish person and number agreement morphology comes from a study by Bruhn de Garavito (2003a,b), who shows that subjects substitute one verb ending for another.[21] Meisel (1991) similarly reports persistent variability in the production of verb agreement in L2 German, with frequent substitution errors for a few speakers. Thus verbal morphology in general is subject to variability that does not necessarily involve presence versus absence of overt markings.

Incorrect inflection surfaces in nominal morphology as well. L2 Spanish gender gives another clear example of the inappropriate use of nominal inflection:

15. la barba rojo (White et al 2004:119)
the.FEM beard red-MASC
'the red beard'

In this example, masculine *–o* agreement surfaces in the adjective *rojo,* replacing the target *roja* (red-FEM); this is not a case of missing inflection, since the adjective is indeed inflected. Similar cases of masculine inflection in feminine contexts in L2 Spanish are reported in Franceschina (2001) and White et al (2004) for determiners

[21] I discuss this study in further detail in Chapter Four, which deals with L2 Spanish verbal morphology.

and adjectives.[22] For both these studies, accuracy with masculine agreement was much higher than with feminine agreement. Although the source of errors in gender agreement is controversial, the type of errors that learners make is not: the majority of errors in L2 Spanish gender involve masculine gender in feminine contexts.[23]

To sum, systematicity cannot be described as a lack of overt morphology, or, put differently, the overuse of default zeros. In this section, several cases that involved an alternation between two overt markings were cited, both for verbal and nominal morphology. Nevertheless, whether the alternation is between zero and overt forms, or between two overt forms, the substitutions seem to be quite systematic.

3.1.3 Task effects

Thus far, I have presented evidence for morphological variability that primarily comes from production data. The issue of variability in comprehension has received less attention. Nevertheless, a few studies allow us to speculate to what extent (if any) variability crosses into comprehension.

The occurrence of morphological variability in production is, in some ways, inconclusive. For some authors, morphological variability is taken to be indicative of a lack of underlying knowledge; variability reflects a lack of competence. For others, especially those supporting the MSIH or similar proposals, variability in production is attributed to problems of lexical access; difficulties arise when the pressure to communicate is strong. Under this view, production data may underestimate the underlying syntactic *competence*; the non-targetlike language that we can observe is a product of difficulties related to *performance*. Montrul (2004), for example, concludes that errors arise because "the mature *performance* system is set in a way that becomes inflexible to accommodate structures that differ from those of the L1" (p. 371; emphasis added).

In principle, then, variability in production may mean either that L2 knowledge is truly non-targetlike, or that it is (more) targetlike but that we cannot gain insight into to this knowledge given the nature of the task. If errors are merely a product of

[22] I revisit this study in greater detail in Chapter Five, which deals with gender morphology.
[23] In L2 French, Hawkins (1998) reports on a study showing that the choice of default may actually vary: some L2ers use feminine as a default, and others use masculine.

performance limitations, then we should be able to observe something closer to the underlying competence by engaging L2ers in some task that minimizes the pressure placed on the learner, perhaps by tapping into comprehension.

In addition to the question of whether variability extends to comprehension, an important question is whether comprehension variability, if it exists, is of a similar nature to production variability. For example, do learners assume the same default forms that they adopt in production? If they do, this might constitute an argument for an underlying representational issue driving the use of default morphology across methodologies. Below I review some studies of variability across tasks, with particular reference to the defaults employed, where this type of data is reported.

I will begin with a discussion of verbal morphology across tasks. White (in press) compares 3rd singular *–s*, past *–ed,* and determiners in L2 English across two tasks, a grammaticality judgment and a production task. The grammaticality judgment task was a preference task, involving the selection of one sentence as grammatical within a pair; one sentence contained incorrect morphology and the other contained correct morphology. The production task involved describing a series of pictures. Two L1 groups were examined: L1 French and L1 Mandarin. For the variable of missing 3rd person *–s*, the French group's accuracy rate was higher in the comprehension task (a mean of about 8 of 9) than in the production task (about 60 percent), suggesting less variability in comprehension. However, the French group still showed variability in comprehension, in that they performed significantly worse than native controls. Thus some evidence for a task effect is found for *–s* usage, though it appears that variability in comprehension had surfaced.

In Bruhn de Garavito's (2003a,b) study of L2 Spanish verbal agreement morphology, learners perform significantly better in a comprehension task than a production task.[24] Nevertheless, a correlation between comprehension and production is found across subjects, which is unexpected under an account that places the source of the problem exclusively in production limitations.

In nominal morphology, Franceschina (2002) examines the comprehension and production of gender in Spanish. Two tasks, one that targeted the production of

[24] I revisit this study in greater detail in Chapter Four, which deals with verbal inflection.

pronouns missing from a sentence, and the other that targeted the comprehension of clitics inflected for gender agreement, showed that L2 Spanish speakers (L1 English) were significantly worse than natives at producing and identifying correct gender inflection. Thus variability in gender appears to extend to comprehension. However, Franceschina does not identify the types of errors or the defaults employed, instead arguing that morphological variability means that English natives cannot acquire gender as a functional feature in L2.[25]

One study that compares gender errors in comprehension and production is that of White et al (2004). This study of L2 Spanish gender and number finds that subjects use masculine gender as a default in comprehension as well as production. Their comprehension task involved the selection of an object that corresponded in gender to a determiner in a null-nominal construction (*el nuevo* 'the new one' in 16). Both L1 English and L1 French subjects performed significantly better in identifying feminine determiners as corresponding to feminine objects than doing the reverse: identifying masculine determiners as corresponding to masculine objects. In other words, masculine determiners like the one in (16) acted as a default in comprehension, as subjects extended them to feminine contexts by choosing a feminine object to correspond to the determiner.

16. ¿Dónde puse el nuevo que compré?

(White et al 2004:126, ex.17b)

Where put-1S the-MASC new-MASC that bought-1S

'Where did I put the new one that I bought?'

Low proficiency groups, in fact, selected masculine determiners as referring to feminine objects at a rate of 50 percent. Accuracy was significantly higher for feminine determiners (61 percent for L1 French, and 75 percent for L1 English). Unexpectedly, subjects' overall performance on gender items on the comprehension

[25] The assumption here, of course, is that a low accuracy rate for the feature of gender entails the inability to acquire both masculine and feminine features. This might be an accurate assumption if both masculine and feminine genders had demonstrated high error rates across comprehension and production. However, if the data involved mainly errors in only one of the gender features, the assumption that the entire category of gender is unacquirable would be unwarranted.

task is significantly worse than on the production task. White et al note that the use of defaults in comprehension counters Prévost and White (2000b)'s suggestion that defaults are confined to production.

White (in press) also reports on the judgments and production of determiners by L1 French and L1 Mandarin speakers. For indefinite determiners, the French group was significantly less accurate than controls in the grammaticality judgment task, with a mean accuracy rate of about 4 out of 5, or 80 percent. In production, however, they were not significantly different from natives. The finding that the French speakers appear to have performed worse in the judgment task is surprising given that the task might be expected to give a better reflection of the underlying competence of the L2 speaker than a production task. The pattern of variability across tasks is complex: although it appears that the non-use of 3rd *–s* may be predominantly production-based, the same does not appear to hold for determiners. Thus the issue of morphological variability across tasks appears to be complicated, although some evidence suggests that it may extend to comprehension and grammaticality judgments.

3.1.4 Summary of generalizations on morphological variability

To sum, studies of morphological variability in verbal and nominal morphology lead to the following preliminary generalizations:

- Variability does not entail syntactic deficits (i.e. problems with word order and Case)
- Variability yields defaults that may be either zero or overt
- Variability extends across proficiency levels
- Variability may extend to comprehension and grammaticality judgments

In order to both rule out performance/communication pressures as the sole cause of variability in production and to rule *in* some sort of representational issue that drives variability across the board, we need to examine whether the patterns of errors in comprehension and production are qualitatively similar. For the most part, this

question has been ignored.[26] If different patterns or defaults arise under different task conditions, this would argue against a representational cause for variability, as the representation—whatever it looks like—should yield consistent effects regardless of the task. In order to rule out a strict performance-based account, therefore, comprehension and production tasks will need to yield the same defaults. For the studies of gender and determiner usage cited above, this appears to be the case.

In the following section, I will introduce a proposal that claims that morphological variability is, at least in part, a representational phenomenon, and therefore predicts variability in both comprehension and production. However, unlike current representational deficit accounts (e.g. the FFFH), the source of variability lies in the morphology, not the syntax.

3.2 The proposal

Under DM, bundles of abstract features[27] are manipulated in the syntax. I further assume that features are monovalent, which means that features are either present or absent and no +/- valuation is necessary (for an argument in favor of monovalency, see Harley 1994). Vocabulary insertion is a competition in which the most highly specified vocabulary item, barring feature clash, is inserted into the fully-specified syntax. Where no features match between the terminal node and feature bundle, an *elsewhere* form is inserted. Competition for vocabulary insertion proceeds from the most highly specified entry to least specified entry (the elsewhere form). Where there is an equal number of features specified for two or more forms, I assume that the order is stipulated (following Halle & Marantz 1993).

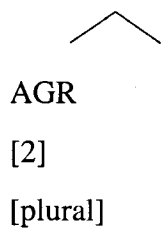
---

[26] A few studies do report on the type of errors across comprehension and production. Bruhn de Garavito (2003a,b) has no comments on the similarity in the type of errors made. White et al (2004) note that the errors in gender are similar across tasks, and this is unexpected under the MSIH. Franceschina (2002) does not report the gender error types.

[27] At this point, I follow Halle and Marantz (1993) in referring to "bundles" of features. The issue of bundled features versus hierarchically-organized features is further addressed by Harley (1994), Bonet (1995), Harley and Ritter (2002), among others. See also Chapter Two and Chapter Six for a discussion of how features are best represented.

In order to illustrate how vocabulary insertion operates, consider the syntactic context of 2nd person plural. In (17), a fully-specified syntactic terminal node[28] (17a) interfaces with the vocabulary items in (17b). Competition for vocabulary insertion begins with the first, most highly-specified vocabulary item *-amos*; this item realizes [1][plural], and since [1] clashes with [2], it is not inserted. The next vocabulary item to be considered is *-an*. This item matches for the feature [pl], but there is neither a match nor a clash with [2]. No feature clash means that *-an* is inserted. The stipulated ordering of *-an* before *-as* ensures that *-an* will always win in the competition for vocabulary insertion when the terminal node is specified for the features [2][plural].

17. a) syntactic terminal node



b) vocabulary items

[1][plural]	↔	-amos
[plural]	↔	-an
[1]	↔	-o
[2]	↔	-as
elsewhere	↔	-a

In accordance with the notion that 3rd person and singular are unmarked, [3] and [singular] are underspecified in (17b): only [1], [2], and [plural] are available as person-number features. If we now consider a syntactic terminal node bearing the features [3][singular], we would find no matching features, only clashing ones, in (17b). The only option for such a terminal node is *-a*, a realization of the Elsewhere Condition.

In principle, two types of errors may occur in the competition for lexical insertion: feature clash and underspecification. Suppose the syntax supplies the [singular] feature, but instead *-an* is produced, a realization of the plural feature:

[28] I am limiting discussion to only person-number agreement features at this point. I also assume part of being fully specified means encoding tense and aspect information and maybe some other information as well. I will put this aside for the moment for the purpose of simplicity.

18. ella hablan
she speak-3PL

This is an error of feature clash between the syntax, which supplies [singular], and the vocabulary item, which is associated with [plural]. Suppose, on the other hand, that the syntax supplies the [plural] feature, but instead *-a* is produced, a realization of the elsewhere condition:

19. ellos habla
they speak-3SG

This is an error of underspecification: where the more highly specified form *-an* should have won the competition for vocabulary insertion, the underspecified form was chosen instead. This does not result in feature clash, since the elsewhere morpheme represents an absence of features.

For the category D, the same logic applies: if the syntax supplies the feature [masculine], the insertion of a feminine form results in feature clash:

20. la libro
the-FEM book-MASC.SG

The insertion of a masculine determiner in a feminine context results in an error of underspecification, as the elsewhere morpheme *el* represents an absence of gender features:

21. el noche
the-MASC night-FEM.SG

The same mechanisms apply for the other variables under consideration: number (both in verbal morphology and nominal morphology), tense, and finiteness. Singular

acts as a default in plural contexts, present (nonpast) as a default in past contexts, and nonfinite as a default in a finite context.

The assumption of underspecification creates an asymmetrical relationship between features: this asymmetry makes the grammar *representationally* different from a grammar under a full-specification theory. Under full specification, all features are represented, and so any asymmetrical behavior observed between features must be derived by stipulation or extra mechanism (such as a feature hierarchy, for example). I will therefore be comparing an underspecification-based grammar to a "naive" full specification grammar—that is, one that accords all features a representation without stipulation as to how features might relate to one another.

The proposed representational difference between features allows for defaults to be predicted, with the following hypothesis in place, which I refer to as the Morphological Underspecification Hypothesis (MUSH):[29]

22. L2 errors are instances of underspecification, not feature clash.

The hypothesis in (22) will be compared to the null hypothesis: that features are represented symmetrically, something that we would assume under a "naive" full-specification theory. Under such an approach, there would be no representational difference between opposed features, nor any external mechanism in place to derive any difference in behavior between opposed features. In practice, for L2, the null hypothesis predicts that there will be no difference between accuracy rates for opposed pairs of features (e.g. masculine and feminine). If the null hypothesis is supported, we expect to see no systematic defaults attested in either comprehension or production; that is, errors of substitution should be random.

In addition to testing the predictions of the MUSH against the null hypothesis, I will also discuss the results of the experiments in light of the predictions of four other theories that attempt to explain morphological variability: the MSIH, the FFFH, and

[29] In previous incarnations (e.g. McCarthy 2006), this was abbreviated as MUH. (The S here stands for specification.)

Lardiere's feature assembly hypothesis. I will elaborate on the predictions of these three approaches in the following section, and show how their predictions differ from those of the MUSH.

3.3 A comparison of predictions of four accounts of variability

I will begin this section by elaborating the predictions of the MUSH in greater detail. The predictions of the MUSH differ from the three theories I discussed above, the MSIH (Prévost & White 2000b), the FFFH (Hawkins & Chan 1997), and the feature assembly hypothesis (Lardiere 1998a,b, 2000, 2005). I will discuss the predictions of each of these below, and show how they differ from the predictions of the MUSH in terms of what types of errors are predicted, and what kind of task effects we might expect with reference to comprehension and production.

Consider the feature specifications in Table 3. Assuming no errors, vocabulary insertion dictates that *hablabas* surfaces in a syntactic context specified for person, tense, and finiteness; *hablaba* surfaces in a context specified for tense and finiteness; *habla* surfaces in a context specified for finiteness; finally, *hablar* surfaces in a context that is underspecified for all features.

Table 3. Feature specifications corresponding to vocabulary items

	hablabas	***hablaba***	***habla***	***hablar***
finiteness	[finite]	[finite]	[finite]	[ø]
tense	[past]	[past]	[ø]	[ø]
person	[2]	[ø]	[ø]	[ø]

The hierarchies in (23) further illustrate the same point: the MUSH predicts that there will be no feature clash, which allows for both (totally) underspecified morphology and *less* specified morphology to act as a default. Thus *hablar* is less

specified than *habla*, which is less specified than *hablaba*, which is less specified than *hablabas*. An L2 learner who commits an error could potentially insert *hablar* in any syntactic context, since it is the least specified of any of these verbs. As long as a particular vocabulary item is less specified than the target, it is fair game for vocabulary insertion under my proposal.

On the assumption that nonfinite forms are underspecified for finite (as well as for person and number features), they can act as a default in any finite context. 3rd singular can act as a default in 1st or 2nd contexts (singular or plural), as well as 3rd plural contexts. 3rd plural, although specified for the feature [plural], can in principle surface as a default in a 1st plural context, as it is underspecified for person. It is predicted that any given item can replace an item to its right, but not an item to its left; items that appear in the same position in the hierarchy cannot replace one another. (The symbol ">>" should be read as "can replace"—the item to the left of the symbol can replace the item to the right.)

23. a) infinitive >> 3rd singular >> 3rd plural >> 1st plural/2nd plural
    b) infinitive >> 3rd singular >> 1st singular/2nd singular >> 1st plural/2nd plural

For tense, the following hierarchy is derived:

24. a) infinitive >> present (nonpast) >> past

Similarly, a masculine singular, masculine plural, or feminine singular determiner could emerge as an underspecification error when the syntactic context of D is [feminine, plural]. A masculine plural cannot act as a default in a feminine singular context, nor vice versa, since either of these contexts would result in a feature clash. These predictions are summarized in the hierarchies in (25a,b) below.

25. a) masculine singular >> feminine singular >> feminine plural
    b) masculine singular >> masculine plural >> feminine plural

On a related note, the MUSH also makes predictions about the distribution of defaults in syntax: no variability is predicted in the syntactic context corresponding to the least marked form. To illustrate, consider a syntactic context of a nonfinite verb: the fully-specified syntax contains the specification [nonfinite]. Returning to the feature representations in Table 3, the only possible option for vocabulary insertion that avoids feature clash is the infinitive *hablar*. The insertion of *habla*, for example, yields a feature clash, since *habla* is specified for [finite]. Therefore, in the least marked syntactic context, no variability is expected.

Extending this logic to the nominal domain, for gender and number, the least marked syntactic context (masculine and singular) is predicted to show no variability under the MUSH. In such a context, the syntax would bear the specification [masculine, singular]; the insertion of any form other than masculine and singular (which is totally underspecified) would yield a feature clash. To summarize, in addition to predicting what types of errors are allowed, the MUSH predicts that no variability occurs in the least marked syntactic context.

The MUSH, by assuming underspecified representations, predicts that variability may surface in both comprehension and production. This prediction derives from the fact that the same representations should be accessible whether we are dealing with comprehension or production. [30] Furthermore, since the same (underspecified) representations are available across comprehension and production, we should expect that variability should be qualitatively similar: that is, underspecified morphemes should act as defaults across both domains. As we will see, comprehension is an important point on which the MUSH and MSIH differ, but on which the MUSH and FFFH agree.

[30] The process of language comprehension is generally not dealt with in formal generative linguistics (e.g. Chomsky 1995), which concerns itself only with competence, and generally speaking, production rather than comprehension. Models of comprehension generally fall under the scope of psycholinguistics. I will not provide a review of models of comprehension, but instead note that I assume the conditions for vocabulary insertion (rules of exponence) to be invariant: it should not matter how these conditions for vocabulary insertion are accessed—whether in comprehension or production. This means that the underspecification we find in the rule of exponence corresponding to a masculine clitic will in principle be accessible whether we access data from comprehension or production (see Section 2.2.2 for a discussion of rules of exponence and underspecification).

3.3.1 The MSIH

The MSIH predicts that when inflection is supplied, it is accurate; this is predicted because features and feature-checking mechanisms are intact. Therefore, aside from finiteness (that is, the non-suppliance of inflection), the MSIH makes no predictions regarding person, number, gender[31], and tense in verbal inflection, only to say that tense and agreement should be accurate if they appear. The MSIH and MUSH make the same predictions regarding finiteness: under both theories, nonfinite forms may act as defaults in finite contexts, but the reverse may not occur. Both theories attribute the default status of nonfinite forms to the issue of underspecification.

Regarding the syntactic context of variability, the MSIH, like the MUSH, predicts no variability in nonfinite contexts. That is, only nonfinite verbs may appear in nonfinite contexts. In finite contexts, variability may only involve the substitution of nonfinite verbs. This contrasts with the MUSH, in which we may find substitutions between finite forms, so long as feature clash is avoided.

Data from comprehension may crucially distinguish the MSIH from the MUSH. Under the MSIH, comprehension deficits are not expected, as errors arise due to communication pressures during production.[32] The MUSH, on the other hand, proposes that there are representational issues[33] behind default morphology, and so comprehension deficits may surface alongside production deficits.

3.3.2 The FFFH

The predictions of the MUSH can also be contrasted with those of the FFFH. The FFFH predicts persistent variability as a consequence of the unavailability of features. Features are unavailable, and thus there is no reason to expect any asymmetrical relationship to emerge: if gender is not available, there is no principled reason to

---

[31] In its initial formulation, the MSIH was not specifically applied to the variable of gender, though White et al (2004) suggest that the use of masculine defaults may be a consequence of underspecification.

[32] It is not totally clear that Prévost and White (2000b) intend for "communication pressures" to be an explanation for the existence of variability in production. Nevertheless, White et al (2004) note that comprehension variability is unexpected under the MSIH. I will therefore take the "communication pressures" explanation seriously, but acknowledge that perhaps it was not intended as a full explanation.

[33] These issues are not to be confused with representational deficits: the underspecification of features is a property of native grammars, and so cannot be considered a deficit.

expect masculine to emerge as a default, for example. Thus the FFFH makes similar predictions to the null hypothesis with respect to feature asymmetries, or at least in the cases like gender, where the L1 did not represent the feature (see Section 3.1). Regarding the syntactic contexts of variability, the FFFH predicts variability to occur anywhere the L2er attempts to represent the feature in question. Contra the MUSH, we should not see any difference between masculine and feminine contexts, for example, since the problem lies in the representation of gender on the whole, not masculine or feminine in particular.

The FFFH would also appear to predict no effect across tasks, since the same impaired representation drives behavior across comprehension and production. Thus we would expect variability to appear in both comprehension and production.

3.3.3 Feature assembly

The predictions of Lardiere's feature assembly hypothesis (Lardiere 1998a,b, 2000, 2005) are similar in principle to those of the MUSH. Lardiere notes, in elaborating the notion of *morphological competence* as an explanation for morphological variability:

> I don't mean performance issues, like problems with lexical retrieval, automaticity, or online performance pressure due to something like 'cognitive overload' although I do think these may well contribute to morphological error in production. Morphological competence includes, most obviously, the knowledge of which forms 'go with' which features. [...] In which domains are various features expressed, in combination with what other features, and why is supposedly the same feature expressed in some domains in some languages but not others? (2005:179)

To Lardiere, the issue behind morphological variability is not simply one of recognizing that a language represents gender or number, for example, but rather involves the complex process of reassembling features from the way that they are bundled and realized (or not realized) in the L1 into the ways that they are bundled

and realized (or not realized) in the L2. The feature assembly hypothesis, then, rejects a central claim of the FFFH: that the lack of representation of a given feature means that L2ers will not be able to represent this feature. From Lardiere's viewpoint, this claim would appear to oversimplify the problem.

While the feature mapping hypothesis provides an important insight into why morphological variability might arise, it does not offer predictions for *which* features might be implicated in variability, nor does it predict asymmetrical relations. Indeed, Lardiere notes (2005:190) that the feature assembly hypothesis is still speculative and remains to be elaborated. Nevertheless, by placing the source of variability in the way that features are represented, not in the way that they are accessed, the feature assembly hypothesis would appear to predict variability across comprehension and production, along the same lines as I propose under the MUSH.

3.3.4 Summary of predictions

To summarize, the predictions of the MUSH, MSIH, FFFH, and Feature Assembly Hypothesis are presented in Table 4. In this table, 'possibly' means that the occurrence of a given characteristic is not excluded; for example, variability in production is *possible* under the MUSH and MSIH, but it is not necessarily *expected* (cf. 'yes' for the FFFH, where variability in production *is* necessarily expected, as deficits implicate variability).

Table 4. Summary of predictions for four accounts of variability

	Variability in production	**Variability in comprehension**	**Feature asymmetries**
MUSH	Possibly	Possibly	Yes
MSIH	Possibly	No	Possibly (finiteness: yes)
FFFH	Yes (random)	Yes (random)	No
Feature Assembly	Possibly	Possibly	Possibly

3.4 Conclusion: Variability and default morphology

Morphological variability is a robust phenomenon attested across a variety of L2s. Research thus far has focused on what the syntactic consequences of variability might be, rather than the specific features that might be involved when alternations are attested. While variability primarily involves the presence versus absence of morphology in L2 English, the same cannot be said for Spanish, where we find variability that involves the alternation between two overt affixes. Furthermore, variability may extend into comprehension, though only a few studies have dealt directly with this point. Comprehension is a crucial testing ground because it will allow us to determine whether variability is strictly a performance-based issue that is limited to production, or whether deeper representational issues are driving linguistic behavior, and are consequently evidenced across methodologies.

In this chapter I also presented the MUSH, which claims that L2 errors are predictable. This contrasts with the three other proposals, the MSIH, FFFH, and feature assembly hypothesis; these alternative proposals, as they stand, offer no predictions regarding the kinds of substitution errors that might occur (excluding the MSIH for finiteness). The MUSH predicts that L2 errors are ones of underspecification, rather than feature clash. Assuming the feature representations established in Chapter Two, the MUSH predicts that masculine, singular, 3rd person, present, and nonfinite will surface as defaults in feminine, plural, non-3rd, past, and finite contexts respectively—rather than the reverse. Thus the MUSH is, at least in part, a representational account that predicts variability across comprehension and production. The predictions of this theory will be tested in the following two chapters: Chapter Four for verbal morphology, and Chapter Five for nominal morphology.

Chapter 4
A study of underspecified inflection in the verbal domain

This chapter addresses the acquisition of L2 Spanish inflection in the verbal domain—specifically, the features of person and number agreement, tense, and finiteness. As previously discussed in Chapter Three, there has been considerable debate on the syntactic consequences of morphological variability: when variability occurs, does it entail a syntactic deficit (e.g. Hawkins & Chan 1997)? Evidence indicates that there is a dissociation between morphology and syntax: while syntactic diagnostics such as case assignment and verb raising suggest that learners do have functional projections and feature strength associated with case and raising, morphological variability persists in the form of errors in tense and agreement (see Lardiere 1998a,b, and Chapter Three of this dissertation for further discussion). This chapter looks closely at another aspect of variability in the verbal domain: the features implicated in default morphology. The results of two experiments will be reported here: one on comprehension, and one on production.

Previous research on the acquisition of tense, agreement, and finiteness in the generative framework has focused primarily on L2ers' accuracy in production. As discussed in Chapter Three, most studies come to the conclusion that when inflection is present, it is accurate: the real problem, therefore, appears to be in supplying inflection (e.g. White 2003). Evidence to support this generalization comes from the fact that L2ers generally do not oversupply inflection—that is, they do not produce tense and agreement morphology in contexts in which it does not belong.

Morphological variability in L2 Spanish presents a somewhat different picture. Two studies of L2 Spanish verbal morphology (Mezzano 2003, Bruhn de Garavito 2003a,b) show that overall, L2ers are highly accurate with person and number agreement. These studies nevertheless cast doubt on the "if present, then accurate" generalization that I described in Chapter Two by showing that errors in inflection frequently involve alternations between two morphologically-overt, finite forms.

Mezzano (2003), in a study of four beginning L2ers in a set of elicited production experiments, finds a very low rate of errors (around 4 percent). Nevertheless, a wide range of errors are attested: finite-for-finite substitutions as in (1), an error of number; nonfinite-for-finite substitutions (2); and even finite-for-nonfinite substitutions (3).

1. Los hombres que **juega** con *fire.* (Mezzano 2003:16)
   the men that play-3SG with fire
   'The men that play with fire.'
2. La chica **hablar** con el chico (Mezzano 2003:15)
   the girl speak-INF with the boy
   'The girl speaks with the boy.'
3. ¿Te gusta **salas** comer? (Mezzano 2003: 14)
   you like go-out.2SG eat-INF
   Do you like to eat out?

Across two taping sessions, the percentage of errors involving finite verbs in non-finite contexts decreased, while the percentage of errors involving non-finite verbs in finite contexts increased (see Montrul 2004 for a discussion).[34] As for person, randomness in errors is also reported by Mezzano: the percent of errors that involved 3rd person defaults was 50 percent in the first session, and 57 percent in the second session. Subjects therefore display a preference, though perhaps a weak one, toward the overuse of 3rd person in non-3rd contexts: strictly speaking, true randomness would equal 3rd person defaults one-third of the time (and 1st person defaults one-third of the time, and 2nd person defaults one-third of the time). Mezzano's findings pose a problem for the Missing Surface Inflection Hypothesis (MSIH; Prévost & White 2000b); recall that the MSIH predicts that nonfinite forms may substitute for finite ones, but that finite-for-finite substitutions should not occur. Her findings also pose an immediate problem for the MUSH, as many of the person errors do not

[34] A comparison of Mezzano's results with those reported here will show that errors in finiteness appear to be short-lived, in that they do not extend to L2ers at higher levels of proficiency.

conform to the predicted pattern of 3rd person defaults; I will return to this in Section 4.2.3.

In a study of person and number inflection in beginning-level Spanish, Bruhn de Garavito (2003a,b) finds errors in person agreement across comprehension and production, though once again L2ers are quite accurate overall. In an elicited production task, subjects made person errors at a rate of 10.1 percent. 66.3 percent of errors involved the use of 3rd person for another person, 13.5 percent the use of an infinitive in place of a finite verb, 12.4 percent 1st person for another person, and 7.9 percent 2nd person for another person. Bruhn de Garavito attributes the overuse of 3rd singular to the fact that the infinitive bears a suffix (*-r* as in *hablar*), whereas 3rd singular bears only a stem and theme vowel, as in *habla* 'he/she/it speaks' (2003b:19).[35] Bruhn de Garavito therefore appears to conclude that the L2ers choose the default that bears the least amount of (overt) inflectional morphology, and in the case of regular Spanish verbs, this is 3rd singular. According to this explanation, the problem lies in the use of overt morphology and is not necessarily related to the features that correspond to the overt morphology.

However, an explanation based on the quantity of overt morphology falls short. A missing *-r* accounts for many of the errors of regular inflection, or about 59 percent of errors, but it does not account for the most frequent single error in the data set. This comes in the use of the verb *ser*, which makes up 23.6 percent of total errors. Of these, 17 of 21 involve the use of *es* 'is' in place of *eres* '(you) are'. For this error a missing affix is clearly not the problem, as no omitted affix[36] can yield *es* from *eres*. On the other hand, the substitution of 3rd person *es* for 2nd person *eres* suggests that features, rather than affixes, may be the source of the problem. In addition, a missing

[35] Although Bruhn de Garavito does not make the suggestion, the 3rd singular defaults that she finds could be analyzed as nonfinite, making the results consistent with the MSIH, which proposes that nonfinite verbs may surface as defaults in finite contexts, but that finite verbs should not exhibit variability. The regular 3rd person verbs that lack *–r* and surface in 1st and 2nd person contexts could be plausibly analyzed as nonfinite, but it seems unlikely that *es* 'is' is a nonfinite verb given its lack of resemblance to the infinitive *ser* 'be'. A feature-based analysis would analyze both the *es* cases and the regular 3rd person cases as a unitary set of phenomena involving the substitution of 3rd person for 1st/2nd. Bruhn de Garavito, however, suggests (p.c.) that *es* may in fact be a nonfinite verb, if we suppose that *s-* is the stem, and that *e-* is epenthesized in order to construct the phonologically-plausible word *es*.

[36] Spanish only uses suffixes in inflection (see Chapter Two for a description). I assume that the learner has not mistakenly omitted a prefix *er-* from *eres* to yield *es*.

–r would seem to predict that defaults would maintain the stem shape of the nonfinite verb. Several examples I will site below suggest that defaults do not maintain the nonfinite stem, but rather take on the finite stem allomorph. In (10) below, *puede* (infinitive *poder* 'be able to') surfaces as a default; a missing *–r* would incorrectly predict *pode*.

To measure comprehension, Bruhn de Garavito administered a multiple-choice task which consisted of selecting the correct subject of an agreeing verb. Subjects performed slightly better on this task, with an error rate of about 4.5 percent. Under a theory that predicts variability to be a production-based phenomenon, such as the MSIH (Prévost & White 2000b) (see also Chapter Three), any errors might be problematic, but the error rate is so low that it is almost negligible. Overall, variability appears in both comprehension and production of person and number agreement, but the amount of variability is relatively small under both comprehension and production. Nevertheless, the quality of errors that do surface reveal something about the representation of features and the use of default morphology.

In the acquisition of tense morphology in L2 Spanish, most studies have looked at the acquisition of the preterite-imperfect aspectual contrast (e.g. Montrul & Slabakova 2002) rather than contrast between past and nonpast/present tense. Nevertheless, Mezzano (2003) reports that beginning L2 Spanish speakers do produce errors in tense, using present-tense verbs in past and future contexts, as in (4).

4. En el pasado, me gusta mucho mi trabajo. (Mezzano 2003:18)
   In the past, to-me like-PRES a lot my job
   'In the past, I liked my job a lot.'

In this example, the target form would have been the preterite *gustó* or imperfect *gustaba*. The L2ers were highly accurate in present-tense contexts, that is, they did not use past-tense morphology in contexts where present would have been appropriate. Importantly, the low-proficiency L2ers in Mezzano's study had not been taught the past or future at the point of their interviews, and so the overuse of present-tense as a substitute is unsurprising. The use of tense by L2ers at higher levels of

proficiency, who have been exposed to the past tenses of Spanish, is an area that merits further investigation, and one that I will explore here.

To conclude this brief summary of research in L2 Spanish verbal morphology, it is clear that morphological variability needs to be investigated at the level of features. In light of the fact that missing affixes do not tell the whole story, a difficulty with morphological overtness or affixation cannot be relied upon to fully explain morphological variability. The hypothesis I will explore in this chapter concerns the representation of features in determining the outcome of morphological variability.

This chapter is organized as follows: Section 4.1 states the predictions for variability in verbal morphology. Section 4.2 reports on a study of spontaneous production in the verbal domain. Section 4.3 reports on a study of comprehension of verbal agreement morphology. Section 4.4 concludes this chapter with a summary of the major findings and a discussion theoretical significance of the results.

4.1 Features in the verbal domain

Chapter Two established the features that I assume, based on markedness relations, to be underspecified in the grammar for the categories of person, number, tense, and finiteness. The underspecified features are assumed to be unavailable in the lexicon, and these terms are represented via the absence of a feature. The feature categories and the underspecified features are listed in (5) below:

5.

Category	*Underspecified feature*	*Other features*
Person	[3]	[1], [2]
Number	[singular]	[plural]
Tense	[nonpast]	[past]
Finiteness	[nonfinite]	[finite]

The specific hypothesis to be tested, repeated from Chapter Three, is the following:

6. L2 errors are ones of underspecification, not feature clash.

The following predictions are therefore made with (6) and the assumed feature specifications in place:

7. a) Person: 3rd person morphology will occur in 1st and 2nd syntactic contexts, but not the reverse;
   b) Number: singular morphology will occur in plural syntactic contexts, but not the reverse;
   c) Tense: present-tense morphology will occur in past syntactic contexts, but not the reverse;
   d) Finiteness: nonfinite morphology will occur in finite syntactic contexts, but not the reverse

These predictions will be tested for both comprehension and production. For production, an underspecification error would involve the substitution of an underspecified form in a context in which the more highly specified form should have occurred. This means 3rd person, singular, present, and nonfinite are predicted to be produced as defaults in 1st/2nd person, plural, past, and finite contexts, respectively.

For comprehension, an underspecification error would involve the interpretation of underspecified morphology as appropriate in a syntactic context that corresponds to the marked form. For example, "3rd person occurring in 1st and 2nd syntactic contexts" will be measured in comprehension as selecting a verb inflected for 3rd person in a 1st or 2nd person syntactic context. Thus, when errors occur, 3rd person and singular morphology are predicted to be comprehended as corresponding to 1st/2nd person and plural contexts, respectively. The comprehension task tests only person and number agreement, not tense or finiteness.

Across comprehension and production, the results of these experiments will be compared to the null hypothesis, which supposes that features are symmetrically represented, and therefore predicts that errors should be random. This chapter also

considers an alternative hypothesis[37] on morphological variability, the Missing Surface Inflection Hypothesis (MSIH), as proposed by Prévost and White (2000b) (see Chapter Three). The MSIH places special importance on finiteness. Under the MSIH, variability may involve an alternation between nonfinite and finite verbs (with nonfinite verbs replacing finite ones), but it may not involve an alternation between two finite verbs. That is, the MSIH predicts that person, number, and tense errors will not occur, but that errors of finiteness will occur. The predictions of the MSIH can be contrasted with those of the MUSH, which accords no special importance to finiteness: any verbal feature may exhibit variability, so long as the errors that arise involve the substitution of underspecified inflection. As we have already seen, data from previous research on L2 Spanish (e.g. Bruhn de Garavito 2003a,b for person agreement, Mezzano 2003 for person and tense) suggests that variability between finite forms does appear; this chapter looks for an explanation in terms of morphological features.

4.2 Spontaneous production

This section reports on a study of errors in the spontaneous production of person, number, tense, and finiteness in L2 Spanish verbs (previously published, in part, in McCarthy 2006). The methodology and results are described below.

4.2.1 Method

The data come from interviews with speakers of L2 Spanish. Eleven participants, all of whom began learning Spanish after age 12, are included in the data set. All participants were asked to rate their level of proficiency in spoken Spanish. Responses ranged from intermediate to advanced, except for one who reported being near-native. Following the methodology of White et al (2004)'s study of grammatical gender in L2 Spanish, proficiency was measured by combining the scores on (1) a

[37] I will not be comparing the predictions of the MUSH to those of the FFFH in this chapter. The FFFH predicts that features will "fail" when they are new to the L2er (i.e. not present and overtly marked in the L1; see Hawkins 2001, White in press for further discussion). Person and number are overtly represented in L1 English (through 3rd singular *–s*), as is tense and presumably finiteness as well (past *-ed*). Thus the FFFH would appear to predict success in these domains as L2ers can represent all of these features.

cloze test from the Diploma de Español como Lengua Extranjera (Spanish Embassy, Washington DC), and (2) a multiple choice test from the reading/vocabulary section of the MLA Cooperative Foreign Language Test (Educational Testing Service, Princeton, NJ). Only those subjects that scored in the intermediate and advanced range are included in the sample. Of these eleven, 6 scored in the intermediate range and 5 in the advanced range. The majority (nine of eleven) of participants had received both naturalistic and classroom exposure to Spanish: they reported having lived in a Spanish-speaking environment for four weeks or more. All had received at least one semester of formal instruction. Four subjects were interviewed, but their data were excluded: one subject was excluded since his data yielded no errors in any of the variables of interest, and the other three were excluded due to French exposure during the critical period. Since the research was conducted in the primarily French-speaking city of Montreal, it was difficult to find true cases of Spanish as a second (not third) language. It was decided that pre-critical period exposure to a Romance language might introduce unwanted variables, since French has similar properties to the L2 target language, Spanish.

Interviews were conducted by a native speaker of Spanish. Participants were told that they should consider the interview a casual conversation, and were encouraged to ask questions of the interviewer if they wanted. The interviewer had a predetermined set of topics and questions, but she was encouraged to allow the participants to talk about any topic that interested them in order to elicit the most naturalistic speech possible. Topics of discussion frequently included travel abroad, academics, family, and daily life in Montreal. Interview lengths ranged from 15 to 35 minutes.

Speech was transcribed by a near-native speaker of Spanish. Following the methodology of Lardiere (1998a,b), utterances that were followed by self-correction were excluded; the final, corrected forms were included. Self-repetitions and repetitions of the interviewer were excluded.

All verbs were coded for person and number agreement, as well as for tense (past vs. present) and finiteness. Errors in person, number, past/present tense, and finiteness were coded for whether the produced form constituted an error of feature clash or

underspecification based on the underspecification of [3], [singular], [present] and [nonfinite].

Data are presented as error counts and accuracy rates. A series of chi-square tests was performed to test for the independence of the factors of feature (non-3rd vs. 3rd; singular vs. plural; present vs. past; nonfinite vs. finite) and accuracy (accurate vs. inaccurate).[38] For some variables, the low error rate made it inappropriate to apply a chi-square test; in these cases, only counts and percentages are given. Statistical significance was set at $p<.05$.

The results for person agreement in verb morphology are broken down into two categories: lexical verbs and copular/auxiliary verbs. This was done because of previous reports in the literature that L2ers are very accurate in the production of agreement and tense in auxiliaries and copula, but not necessarily in lexical verbs (e.g. Lardiere 1998a,b).

4.2.2 Results

Person

A total of 805 lexical main verbs were counted for the analysis of person, only 16 of which involved errors. Accuracy rates for person overall, for non-3rd person agreement, and for 3rd person agreement are given in Table 1. The contingency between person (non-3rd vs. 3rd) and accuracy (accurate vs. inaccurate) is shown in Table 2. All errors in non-3rd agreement involved the substitution of 3rd person agreement in a 1st or 2nd person context; 12 involved the substitution of 3rd for 1st, and 2 the substitution of 3rd for 2nd.

[38] In using chi-square tests I follow methodology of, for example, corpus linguistics, first language acquisition research, and sociolinguistics. However, these tests should be interpreted with an element of caution, as the responses are not truly independent of each other (in the sense that some come from the same speaker). The independence of responses is normally a requirement for using the chi-square test.

Table 1. Percent accuracy rate for person in lexical main verbs for all L2 subjects by category/feature

	Percent Accuracy
Person (all)	98.0
Non-3rd	97.0
3rd	99.4

Table 2. Person agreement in lexical main verbs: feature by accuracy

	Non-3rd	3rd	Totals
Accurate	451	338	789
Inaccurate	14	2	16
Totals	465	340	805

The contingency between person and accuracy is significant (X^2 = 5.92; df = 1, p<.02). It can be seen that there were fewer errors in 3rd person contexts, as predicted by my hypothesis. Errors in 3rd person would necessarily involve feature clash.

A total of 577 copular and auxiliary verbs were coded for analysis. Accuracy rates are given in Table 3. The contingency between person feature and accuracy is significant (X^2 =5.17; df = 1; p<.05; see contingency in Table 4), though the pattern of errors runs counter to the predicted direction (i.e. non-3rd person agreement is more accurate than 3rd person agreement).

Table 3. Percent accuracy rate in copular and auxiliary verbs for all L2 subjects by category/feature

	Percent Accuracy
Person (all)	95.8
Non-3rd	99.2
3rd	94.8

Table 4. Person agreement in copular and auxiliary verbs: feature by accuracy

	non-3rd	3rd	Totals
Accurate	134	419	553
Inaccurate	1	23	24
Totals	135	442	577

This unexpected pattern is due to the behavior of one exceptional subject (Sheila) who uses the same preterite 1st person copular/auxiliary verb (*estuve*) repeatedly in 3rd person contexts. In Tables 5 and 6, her data are excluded. After excluding her data, only one error occurs in copular/auxiliary verbs, as shown in Table 6.

Table 5. Percent accuracy rate in copular/auxiliary verbs for all L2 subjects, excluding Sheila, by category/feature

	Percent Accuracy
Person (all)	99.8
Non-3rd	99.1
3rd	100.0

Table 6. Person agreement in copular/auxiliary verbs, excluding Sheila: feature by accuracy

	non-3rd	3rd	Totals
Accurate	106	361	467
Inaccurate	1	0	1
Totals	107	361	468

(8) and (9) below are examples of errors of underspecification in person agreement. In (8), the participant asks a question that lacks 2nd person singular *-s*.[39] There was a slight pause between repetitions of the verb, as she was waiting for the interviewer to answer her. The interviewer did not understand her the first time, since the verb's agreement indicated 3rd person. (9) was uttered in a context in which the participant was asked for information about herself, making the intended referent 1st person.

[39] An alternative analysis would be that the speaker was using the formal *Usted*, in which case no *–s* was needed. However, since she uses the form *tú* following the verb, this explanation seems unlikely.

8. *Samantha (intermediate L2 Spanish):*

y manejará? (pause) manejará tú?

and drive-FUT.3SG drive-FUT.3SG you

'and will you drive?'

9. *Beth (intermediate L2 Spanish):*

nació en Boston

be.born-PAST-3SG in Boston

'I was born in Boston.'

Number

A total of 806 lexical main verbs[40] and 577 copular/auxiliary verbs were coded for number agreement. 20 of the lexical verbs contained errors, 18 of which involved the use of singular agreement in a plural context (see Table 9). 7 of the copular/auxiliary verbs contained errors, all of which involved the substitution of singular morphology for plural (see Table 10). Accuracy rates for lexical main verbs and copula/auxiliary verbs are given in Tables 7 and 8. Overall, subjects are highly accurate with number agreement, with 99 percent accuracy. These contingencies cannot be supported statistically via chi-square test due to the low expected-cell values.

Table 7. Percent accuracy rate in lexical main verbs for all L2 subjects by category/feature

	Percent accuracy
Number (all)	97.5
Singular	99.7
Plural	89.4

[40] The difference between the number of tokens for person and number is explained by vowel reduction that made it impossible to distinguish between a 3rd person suffix (*-a*) and a 1st person suffix (*-o*).

Table 8. Percent accuracy rate in copular/auxiliary verbs for all L2 subjects by category/feature

	Percent accuracy
Number	98.8
Singular	100
Plural	92.2

Table 9. Number agreement in lexical main verbs: feature by accuracy

	Singular	Plural	Totals
Accurate	634	152	786
Inaccurate	2	18	20
Totals	636	170	806

Table 10. Number agreement in copular/auxiliary verbs: feature by accuracy

	Singular	Plural	Totals
Accurate	485	85	570
Inaccurate	0	7	7
Totals	485	92	577

Examples of number errors are given in (10) and (11).

10. *Linda (advanced L2 Spanish):*

Los italianos puede entender un poco
the Italians can-3SG understand a little
'The Italians can understand a little (Spanish).'

11. *Steve (advanced L2 Spanish):*

hay varias regiones en el norte que me gustó
there.are various regions in the north that 1SG-DAT like-PAST-3SG
'There are various regions in the north that I liked.'

Tense

A total of 1415 lexical main verbs, and 577 copular/auxiliary verbs, were coded for past versus present tense. Out of the main verbs, 274 of these occurred in contexts that were obligatorily past tense. Out of the copular/auxiliary verbs, 136 of these occurred in obligatory past-tense contexts. Subjects were highly accurate with tense, with an overall accuracy rate of 99 percent for both verb types. Accuracy rates are given in Table 11 for lexical main verbs, and Table 13 for copula/auxiliary verbs. Tables 12 and 14 display the contingency between accuracy and finiteness for each verb type. No errors involved the use of past-tense morphology in present contexts, whereas 17 errors involved the use of present morphology in a past-tense obligatory context. The difference between tokens of tense and tokens of person and number stems largely from the fact that existential verbs (e.g. *hay* 'there is/are', *había* 'there was/were') were counted for tense but not person/number, as they do not show person/number agreement. These existential verbs were counted as lexical main verbs.

Table 11. Percent accuracy rate for lexical main verbs: past/present tense by category/feature

	Percent Accuracy
Tense (all past/present)	99.2
Present	100.0
Past	95.6

Table 12. Past/present tense in lexical main verbs: feature by accuracy

	Present	Past	Totals
Accurate	1141	262	1403
Inaccurate	0	12	12
Totals	1141	274	1415

Table 13. Percent accuracy rate for copular/auxiliary verbs: past/present tense

	Percent Accuracy
Tense (all past/present)	99.1
Present	100.0
Past	96.3

Table 14. Contingency table for past/present tense in copular/auxiliary verbs: feature by accuracy

	Present	Past	Totals
Accurate	441	131	572
Inaccurate	0	5	5
Totals	441	136	577

Examples of present-tense morphology in obligatory past contexts are given in (12) and (13):

12. *Tom (Advanced L2 Spanish):*

Cuando nos conocimos, yo no **hablo**
when REFL meet-PAST-1PL I NEG speak-PRES-1SG
ningun palabra en español
NEG word in Spanish
'When we met, I didn't speak a word of Spanish.'

13. *Rachel (Advanced L2 Spanish):*

Nací en la ciudad de Nueva York, pero
be.born-PAST-1SG in the city of New York, but
mis padres se **mudan** en 1985
my parents REFL move-PRES-3PL in 1985
'I was born in the city of New York, but my parents moved in 1985.'

A total of 2000 verbs (copular, auxiliary, and lexical) were coded as finite, and 274 as nonfinite. Unlike the verbs included for tense analysis, finite verbs include the morphological future (e.g. *manejará* 'drive-FUT.3SG from example 5); incidentally,

only four such tokens were found in the entire data set, and all were produced by the same subject.[41] Only four instances of nonfinite forms in finite contexts were found in the data set, and no instances of finite forms in nonfinite contexts were found. Subjects neared 100 percent accuracy in finiteness. Due to the low rate of errors in finiteness, a chi-square test was not performed. An example is given in (14).

Table 15. Percent accuracy rate for all verbs: finiteness

	Percent Accuracy
Finiteness (all)	99.8
Finite	99.8
Nonfinite	100.0

Table 16. Finiteness in all verbs: feature by accuracy

	Finite	**Nonfinite**	**Totals**
Accurate	1996	274	2270
Inaccurate	4	0	4
Totals	2000	274	2274

14. *Beth (Intermediate L2 Spanish):*

yo nunca hacer los platos

I never do-INF the dishes

'I never do the dishes.'

[41] This should not be taken to mean that the L2ers do not know how to use the future tense. Spanish has another future tense that uses the verb *ir* 'go' as an auxiliary in the expression of future time. This future tense occurred very frequently.

From the examples cited above, we can see that errors in person, number, and tense are not confined to L2ers at the lower level of proficiency. All 5 of the advanced-proficiency L2ers made at least one error in person/number agreement (see examples 10, 11 above, and 18 below). This observation constitutes further evidence that, although overall accuracy rates are very high, morphological variability is a persistent problem. Errors in past tense were attested in 4 of the 5 advanced speakers as well (as shown in examples 12 and 13). Finiteness errors (of which there were only two) were, however, confined to the intermediate-proficiency L2ers. Individual results are presented in Appendix A for person, Appendix B for number, and Appendix C for tense.

4.2.3 Discussion

Overall, the results are consistent with the Morphological Underspecification Hypothesis, as proposed in Chapter Three and repeated in (3). This hypothesis, in combination with the feature inventories established in Chapter Two, predicts that 3rd person, singular, present, and nonfinite are employed as defaults in the feature categories of person, number, past/present tense, and finiteness. Across all of these categories, the predictions are supported, barring one exceptional participant for the feature of person; her productions will be dealt with below. Furthermore, for all of these feature categories, the null hypothesis—that features are symmetrically represented— is not supported.

For person and number, the rate of errors reveals that subjects are highly accurate in producing agreement; not surprisingly, they are more accurate than the low-proficiency subjects reported in Bruhn de Garavito (2003a,b). Furthermore, the randomness in person errors reported in Mezzano (2003) is likewise not attested. The major generalization found in this data set is that, when errors occur, they involve the systematic substitution of 3rd person for other persons. The data support the hypothesis that errors are ones of underspecification, under the assumption that [3] is underspecified. If, on the other hand, 1st, 2nd, and 3rd person were represented symmetrically, we would have no reason to expect the errors to be unidirectional. Or,

we would need to appeal to some kind of extraneous mechanism, such as a feature hierarchy, that causes 2nd and 1st person features to delete.

One clear counterexample to the predictions made here is Sheila's repeated use of the 1st person preterite *estuve* 'was' (infinitive *estar* 'be') in 3rd person contexts (15). In (15), she uses the wrong copula; *ser* would be used in this context, not *estar*. She has a tendency to use *estuve* anywhere she means *was*. In fact, she never manages to use the target form *estuvo*, suggesting she does not know what the correct form is.

Intermediate L2 Spanish (Sheila):

15. el objetivo estuve ...
    the objective was-1SG
    'the objective was...'

Estuve occurs 22 times in 3rd person contexts, and never in 2nd person contexts. There is also one instance of the verb *hacer* used in the same way, shown in (16):

Intermediate L2 Spanish (Sheila):

16. hay una que mi novio ya hice
    There-is one that my boyfriend already did-1SG
    'There is one that my boyfriend already did.'

Although this is clearly a problem for the MUSH, it should be noted that she uses 3rd person as a default in present tense. In addition, these forms are irregular past preterite forms, and she shows no evidence that 1st person generally acts as a default in the past preterite, or in any particular tense, for that matter. Ideally, this subject's production data should be compared to her own comprehension data in order to better assess her knowledge of person agreement. If she believed that *estuve* was the correct 3rd form, we could easily dismiss these forms as being due to incorrect knowledge of verbal inflection, as they would not be instances of errors, strictly speaking. The issue of knowledge and agreement is addressed further in Chapter Five, which deals with

gender agreement and knowledge of gender assignment. The problem for gender is slightly different, as the subject may not know the gender of the target item; here, the subject knows the intended person, but may not know what the morphological expression of the intended person is. For person, subjects like Sheila may have incorrect or incomplete knowledge of the forms of agreement (see, for example, Herschensohn 2000).

The overall systematicity of person errors reported here is at odds with the near-randomness reported in Mezzano (2003). This is likely an effect of the higher level of proficiency of the learners reported here. Clearly, the generalization seems to be that errors are somewhat random in the early stages of L2 acquisition, but gradually grow more systematic as proficiency increases. There is no principled linguistic reason to expect this outcome under the MUSH. Nor is there reason to expect this outcome under the MSIH, as finite-for-finite substitutions should not occur at all. Likewise, the FFFH does not predict this outcome, as it holds that 'new' features cannot be acquired in the L2, and so the proficiency level of L2ers should not matter at all. I will return to this issue in Chapter Six, where I propose a way to model increased systematicity via the representation of features.

The pattern of number errors is consistent with expectations: singular inflection surfaces as a default in plural contexts. The errors in number that run counter to predictions are given in (17-18). In (17), the plural verb does not agree with the singular *la cultura de centroamerica y de México* 'the culture of Central America and of Mexico'. This is a complex NP, and the problem might actually be that she intended to say *las culturas* 'the cultures' instead of *la cultura* 'the culture'. Or, the plural verb may agree in plural number with the conjoined NP *Centroamérica y de México.*

17. *Advanced L2 Spanish (Rachel):*

la	cultura	de, de	Centroamérica	y	de México	son, existen	pero
the	culture	of, of	Central America	and	of Mexico	are, exist-PL	but

'The culture of Central America and of Mexico are, exist, but...'

The other error of feature clash in number involved the verb *gustar* 'like/please', one of a class of psych verbs that requires an experiencer and a theme (see Montrul 1998 and references therein). The provided agreement is with the experiencer rather than the theme, which is a nonfinite clause, and which should therefore trigger 3rd singular agreement (18):

18. *Advanced L2 Spanish (Annie):*

Los chicos	les	gustan	pegarse	bien
the boys	3PL-DAT	like-3PL	stick	well

'The boys like to stick to you (?)'

This utterance occurred in a description of an unappealing night club and the boys who frequent it. Given this context, it was inferred that it was the boys who were the experiencers — the ones who experienced the liking. (Ultimately, however, it is not totally clear what she intended to mean.) If this interpretation is correct, this error can be alternatively analyzed as a case of agreement with the wrong argument, rather than incorrect agreement. Under this analysis, the speaker has incorrectly produced agreement that corresponds to the experiencer, rather than the theme. In Spanish, dative experiencers are preceded by a "personal a" marker (as in ***a** ella le gustan los gatos* 'she likes cats'; literally 'cats please her'). The lack of a "personal *a*" marker before the dative experiencer *los chicos* suggests that Annie might be treating the experiencer as a "normal" nominative subject that triggers verb agreement. Under this analysis, this error is not an error of feature clash.

Given the fact that very few errors of finiteness occur relative to errors of person, number, and tense, the Missing Surface Inflection Hypothesis is not supported. Under the MSIH, errors are predicted to involve the substitution of nonfinite verbs in finite contexts. Yet finiteness is the category that the L2ers have the least amount of trouble with. Thus for L2 Spanish, when errors occur, they are more likely to involve finite-for-finite substitutions—3rd for 1st person, singular for plural, present for past-- not nonfinite-for-finite ones. Excluding Sheila's problematic data for person, a total of 65 errors were tallied in the data set. Of these, only 4 (6.3 percent) involved the use of a

nonfinite form in a finite context. The remaining errors (93.7 percent) involved finite-for-finite substitutions, the most common of which being the substitution for singular inflection in a plural context. To summarize, default morphology is best described as underspecified inflection, not missing inflection. The breakdown of errors is given in Table 17.

Table 17. Distribution of errors by feature category

	Number	Percent of errors
Finiteness	4	6.3
Tense	17	26.5
Number	27	42.2
Person	16	25

In at least one respect, the methodology of spontaneous production data is desirable. These interviews approximate naturalistic speech in a way that other laboratory methodologies cannot. The speaker is free to use whatever structures he/she wishes to employ, and is free to change the topic of conversation at will. The level of metalinguistic awareness is extremely low, and so subjects are unlikely to be relying on memorized knowledge in their use of verbal agreement. Since person/number agreement is encoded on all finite verbs, the quantity of data is abundant, and no experimental manipulation was needed.

Regarding the person and number errors that were attested, the majority of person errors involved the substitution of 3rd for 1st person, and the entirety of number errors involved the substitution of 3rd singular for 3rd plural. There is nothing in the feature representations I assume that predicts this outcome. In principle, 3rd for 2nd person substitutions and 1st singular for 1st plural substitutions are equally possible. In the context of spontaneous production, the L2ers simply had less occasion to produce 2nd singular, 1st plural, and 2nd plural, and so fewer contexts existed relative to 3rd plural.

Future research might involve designing a task that elicits speech using a more balanced sample of person and number contexts.

At this point, I would like to point out what seems to be a potential problem for the MUSH. A great deal of evidence indicates that learners have trouble producing 3rd singular agreement in L2 English (e.g. *she walk*) (see Chapter Three and references therein). The MUSH predicts that 3rd person singular acts as a default; it might therefore also predict the overuse of 3rd singular morphology in 1st/2nd person contexts (as in *you walks*). Although this is one predicted outcome, it is not the only predicted outcome. Recall that nonfinite verbs are even less specified than 3rd singular ones. In English, nonfinite verbs are homophonous with 1st/2nd person singular verbs. Thus we could analyze the bare verb in a production like *she walk* as a nonfinite verb that lacks specification for finiteness; this analysis is consistent with the MUSH. Of course, if we assume that L2 English speakers are producing an overabundance of nonfinite verbs, this raises the question of why they produce so many nonfinite verbs in finite contexts, and why L2 Spanish speakers produce so few. The answer may lie in the fact that Spanish L2ers recognize from very early on that Spanish verbs need to bear overt inflection. This point has also been suggested by Bruhn de Garavito (2003a,b).[42]

As far as tense is concerned, I avoided classifying productions as errors if the context did not seem to be truly an obligatory past-tense one. Examples (12) and (13) are good examples of such contexts. There were, of course, much more subtle cases; these were judged as either errors in past tense, or as correct productions of present tense, by consensus between two judges (one a near-native L2 Spanish speaker, the other a Spanish NS). Of course, subjects often switched from the so-called "historical present" to the past while they were describing events, and these switched passages were not counted as errors. The use of the present to signal past time is something that NSs do as well (see example 27c in Chapter Two). I see no way around the difficulty in identifying what is, or is not, an obligatory context for past tense. Whether productions are classified as errors or not, the use of the present to narrate

[42] Of course, a formal linguistic explanation is lacking here, and I leave this as a point to be worked out in future research. It might be the case that a phonological account can explain the missing *–s* in the case of L2 English, as suggested by Goad, White & Steele (2003), but it does not explain the use of finite defaults in L2 Spanish.

past events can be taken as further evidence for the underspecification of [past] in L2 grammars (and perhaps native Spanish grammars as well; see Cowper 2004).

4.3 Comprehension

This section describes an experiment that examines the comprehension of person and number. The goal of this experiment is to provide a complement to the spontaneous and elicited production experiments previously discussed in this chapter.

4.3.1 Method

In order to assess the comprehension of verbal person-number agreement, a written multiple-choice task was given to subjects whose level approximated the subjects in the production component of this chapter. This task was similar in methodology to the one employed by Bruhn de Garavito (2003a,b). A total of 9 subjects participated in this experiment. The level of these subjects ranged from intermediate (N= 7) to advanced (N= 2).[43] The proficiency of the advanced L2ers was measured based on the written proficiency test described in Section 4.2. The proficiency of the intermediate L2ers was measured based on their enrollment in an intermediate-level (i.e. second-year) Spanish class at a major Ontario university.

In this task, subjects were required to select the answer that corresponded in person and number to the verb in the sentence. The task consisted of 30 test items, plus 10 fillers that tested for other aspects of grammar beside person-number agreement. Of the test items, 6 tested for each of the following: 1st singular, 2nd singular, 3rd singular, 1st plural, and the syncretic 2nd plural/3rd plural. Half of these verbs were irregular verbs that involved a change in the stem (e.g. *volver* 'return' vs. *vuelve* 'he/she/it returns'), and half of these verbs were regular. Only simple present forms were included.

[43] Both the comprehension and production experiments included a mixture of intermediate and advanced subjects. However, the comprehension experiment included a higher proportion of intermediate-level subjects. I aimed for a greater number of intermediate L2ers because I anticipated that performance would be better on the comprehension task, and I wanted to capture any variability that might exist.

Sample items: Person-number comprehension task

19. _____ limpia el baño.

a) nosotros b) Enrique c) tú d) ellos e) yo

_____ *cleans the bathroom.*

a) we b) Enrique c) you d) they e) I

20. _____ estoy feliz porque hace buen tiempo.

a) yo b) tú c) María d) María y yo d) ellos

_____ *am happy because the weather is nice.*

a) I b) you c) Maria d) Maria and I e) they

In (19), the correct answer based on the agreement of *limpia* 'clean-3SG' is (b), Enrique. In (20), the correct answer is (a), *yo*, as it agrees with *estoy* 'am'. The test materials are provided in Appendix H.

4.3.2 Results

The results show a lack morphological variability in person and number agreement, as the overall accuracy rate is 99.3 percent (268 out of 270 test items). Out of the 9 subjects, 7 scored 100 percent on the test items. The other two subjects, who were both at the intermediate level of proficiency, made one error each. One was an error involving the selection of the 3rd person pronoun *ella* in a 2nd person context; this is an underspecification error. The other was an error involving the selection of the 1st person plural pronoun *nosotros* in a 2nd person singular context. This is an error of feature clash, as a plural pronoun was selected in a singular context. Because of the overwhelming accuracy in person and number comprehension, these errors cannot be taken as representative of any kind of pattern.

4.3.3 Discussion

The results of the comprehension task do not yield any insight into the inventory of L2 features, as the subjects performed at native-like levels of accuracy. Therefore we do not find evidence to support the MUSH in the comprehension of verbal

morphology; the MUSH predicts that those errors that occur will be of a certain type, and here we have no errors. Therefore the MUSH is neither supported nor refuted.

The comprehension results could be interpreted as evidence for a divergence between comprehension and production, consistent with the claims of the MSIH, which predicts that comprehension should be essentially devoid of variability. However, the MSIH was not supported based on the production data, which showed finite-for-finite substitution errors, inconsistent with the MSIH. Taken together, the comprehension and production results suggest that underspecification applies in the realm of person and number agreement, as evidenced by the production data; however, we have failed to find evidence for underspecification in comprehension. Given the high degree of accuracy in production of a comparable set of subjects, the lack of errors in comprehension is not surprising. Some suggestions for error elicitation in future research are made in Section 4.4.

4.4 Conclusion

The results of the two experiments on verbal morphology support the main hypothesis of this dissertation, which holds that errors involve the overuse of underspecified inflection. In data collected from spontaneous production, it was shown that errors were largely systematic: 3rd person substituted for 1st and 2nd person, singular substituted for plural, present substituted for past, and, though infrequent, nonfinite substituted for finite. While the results are consistent with the predictions of the MUSH, they do not support the MSIH, which predicts no variability between finite forms. Here, we find that a majority of errors involve finite defaults rather than nonfinite ones. For all of the variables examined here, the morphological defaults in production correspond to underspecified inflection. Features are clearly important in explaining morphological variability, as both the MUSH and Lardiere's feature assembly hypothesis claim—the latter, however, does not offer an explanation for why asymmetrical relationships should obtain.

In a comprehension task targeting person and number agreement, essentially no errors were found. This was interpreted as an absence of evidence for underspecification in comprehension.

One question that arises is why learners are so accurate with person and number agreement. The answer might lie in the fact that the subjects were all, to a greater or lesser extent, classroom L2 learners, and L2 Spanish classes typically target verb agreement paradigms from the very first days of instruction. Perhaps by looking at naturalistic learners we might find more informative and robust patterns of errors; simply looking at lower-proficiency learners does not seem to guarantee a high degree of variability (Bruhn de Garavito 2003a,b, Mezzano 2003). In addition, by making the task more difficult—perhaps in terms of time limitations, or by tapping listening rather than reading—we might be able to elicit more errors and more informative data.

Chapter 5
A study of underspecified inflection in the comprehension and production of gender and number agreement

This chapter addresses the acquisition of the features of gender and number in the nominal domain—specifically, in clitics, adjectives, and determiners. In Chapter Three, evidence in favor of an underspecification-based approach to inflection was found for verbal features, including person and number, finiteness, and tense; however, underspecification was found to be limited to the production data. This chapter similarly probes both comprehension and production, but unlike the verbal domain, evidence for underspecification is found in both nominal domains.

Previous studies of gender agreement within the generative framework have focused on the issue of whether or not gender is acquirable in an L2, particularly in cases where the L2er comes from an L1 that does not have gender as a feature (Hawkins 1998, Bruhn de Garavito and White 2002, White, Valenzuela, Kozlowska-Macgregor, and Leung 2004, Franceschina 2001, Sabourin 2003). Findings suggest that in general, the instantiation of gender in the L1 may facilitate the acquisition of L2 gender (Sabourin 2003), although the L1 instantiation of gender is neither a necessary nor sufficient factor in complete L2 success, as L2 gender is sometimes native-like even when not instantiated in the L1 (White et al 2004), and is sometimes problematic even when the L1 has gender (Bruhn de Garavito & White 2002).

This chapter will address the acquisition of L2 gender as a new feature only indirectly; instead, I will focus on the acquisition of the features of masculine and feminine that make up the larger category of gender. The experiments in this chapter will show that, in both comprehension and production, masculine inflection surfaces as a default in feminine contexts. In general, the use of a default gender is fairly well documented. Although some reports indicate that L2ers may differ over which gender

they choose as default (e.g. Hawkins 1998 for French), the literature on Spanish is unequivocal, with masculine emerging as a default across syntactic categories (Bruhn de Garavito and White 2002, White et al 2004, Franceschina 2001). However, the use of masculine as a default is not a derivable outcome in theories as they stand (see Chapter 2 for a discussion of default inflection within L2 theory). For Franceschina (2001), in fact, defaults have no status whatsoever in the grammar. Considering the robust nature of default inflection found in previous studies, at least in the realm of L2 production, it seems clear that the ability to explain defaults in theoretical terms is desirable. In Chapter 2, I showed how defaults are derivable at the interface between syntax and morphology; in this chapter, I will show that the predictions regarding these defaults are borne out.

Number is another feature that is subject to underspecification. Based on the diagnostics in Chapter 2, I predict that singular defaults will emerge across domains. Errors in L2 number agreement have been shown to be quite uncommon (White et al 2004); nevertheless, the errors that surface will be analyzed for the presence of default inflection.

This chapter is organized as follows: in Section 5.1, I state the predictions for the experiments presented in this chapter. In Section 5.2 I describe the first experiment, a study that looks at a corpus of spontaneous production data. Section 5.3 describes an elicited production experiment that elicits clitics and adjectives as domains of agreement. Section 5.4 describes a study of the comprehension of agreement in clitics. Section 5.5 summarizes the three experimental tasks and relates these results to issues in second language theory.

5.1 Features in the nominal domain

Chapter Two established the features that I assume to be underspecified in the grammar. For gender, I concluded, based on evidence from Spanish, that [masculine] is unmarked, and by assumption, underspecified. Similarly, I concluded that [singular] is unmarked, and by assumption, underspecified. In other words, [masculine] and [singular] are underspecified in the lexicon/vocabulary. This feature

inventory follows previously-established generalizations, both with respect to universals (Harley and Ritter 2001 for number) and with respect to Spanish in particular (Harris 1991, Bonet 1995; see also Chapter Two of this thesis).

The specific hypothesis to be tested, repeated from Chapter Two, is the following:

1. L2 errors are ones of underspecification, not feature clash.

The following predictions are therefore made according to (1) and with the assumed feature specifications in place:

2. a) Gender: masculine inflection will occur in feminine syntactic contexts, but not the reverse;

   b) Number: singular inflection will occur in plural syntactic contexts, but not the reverse.

For production, "occurring in feminine syntactic contexts" will involve using a masculine determiner with a feminine noun (as in *el casa*), a masculine clitic to refer to a feminine NP (as in *comerlo* 'eat-INF it-MASC', where *lo* refers to *la manzana* 'the-fem apple'), or a masculine adjective to agree with a feminine noun (as in *casa blanco* 'white-MASC house-FEM'). The same logic applies to number: singular inflection "occurring in plural syntactic contexts" will be measured as using a singular determiner with a plural noun, a singular clitic to refer to a plural NP, or a singular adjective to agree with a plural noun.

For comprehension, "occurring in feminine syntactic contexts" will be measured as accepting a morphologically masculine item as grammatical in a feminine syntactic context. For clitics, this will involve accepting a masculine clitic in a context in which a feminine clitic should have occurred. Similarly, "occurring in plural syntactic contexts" will be measured as accepting a singular clitic in a context in which a plural

clitic should have occurred. The comprehension of gender and number will be tested only in clitics.

5.2 Spontaneous production

This section reports on a study on errors in the spontaneous production of gender and number in determiners (previously published in McCarthy 2006). This study establishes a baseline pattern in which masculine and singular determiners are extended to feminine and plural contexts, respectively. The methodology and results are described below.

5.2.1 Method

The methodology of this experiment, including the subjects and proficiency testing, is described in detail in Section 4.2.1.

All errors in gender and number were coded for whether the produced form constituted an error of feature clash or underspecification based on the underspecification of [masculine] and [singular]. Plural determiners occasionally contained a reduced vowel, which made it difficult to determine whether the target form was *los* or *las*, *unos* or *unas*. These tokens were excluded.

Data are presented as error counts and accuracy rates. As before, a series of chi-square tests was performed to test for the independence of the factors of feature (masculine vs. feminine) and accuracy (accurate vs. inaccurate). Statistical significance was set at $p<.05$.

5.2.2 Results

A total of 105 agreement errors were found for determiners out of a total of 867 determiners. Of total errors, only three showed a number error: all of these involved the occurrence of a singular determiner in a plural context, an underspecification error. Since the error rate for number was a fraction of one percent, I will not provide a statistical analysis for number. Accuracy rates are shown for gender overall, and for masculine and feminine contexts, in Table 1. Subjects were highly accurate with agreement in masculine contexts (97 percent accurate), while agreement in feminine contexts was more variable (77 percent accurate), as subjects substituted masculine determiners for feminine ones more often than the reverse.

A significant contingency was found between accuracy and feature, such that subjects were more accurate with agreement in masculine contexts than feminine ones (X^2 = 85.5, df = 1, p<.0001); see Table 2. Of the 102 errors of gender that occurred, 89 (87 percent) involved the substitution of a masculine determiner for a feminine one—an underspecification error.

Table 1. Percent accuracy rate in determiners for all L2 subjects by feature

	Percent Accuracy
Gender (all)	88.2
Masculine	97.3
Feminine	76.9

Table 2. Contingency table for gender agreement in determiners: feature by accuracy

	Masculine	Feminine	Totals
Accurate	468	297	765
Inaccurate	13	89	102
Totals	481	386	867

The determiner error rates reported in Table 1 are quite similar to the ones reported in the production experiment of White et al (2004). There, the accuracy in masculine agreement was around 97 percent, while feminine ranged from 88 percent for intermediate to 96 percent for advanced in contexts without an adjective, and 68 to 99 percent in contexts with an adjective. In this data set, I have not separated the tokens that occur with adjectives from those that do not; it is not surprising, then, that the accuracy rate for feminine agreement falls between the two ranges reported in White et al. The masculine accuracy rate is almost identical in both studies, and both studies find a significant effect of feature, with masculine agreement more accurate than feminine.

Underspecification errors included errors in indefinite (3) and definite (4) determiners.

Samantha (intermed.)

3. **un** mejor educación

   det-INDEF-MASC better education-FEM

David (intermed.)

4. **el** sangre *acadian*

   det-DEF-MASC blood-FEM Acadian

Errors with nouns that follow the "canonical" gender patterns in Spanish (feminine nouns end in *-a*, masculine ones in *-o*) are surprisingly frequent. A clear pattern emerges with respect to these canonical nouns: there are many (about 30) instances of canonical feminine nouns occurring with a masculine determiner (e.g. *un mezcla* 'a mixture', *un palabra* 'a word', *el revista* 'a magazine', *los islas* 'the islands'), but there are only five instances of a canonical masculine noun occurring with a feminine determiner (*una cuarto* 'a room', *una método* 'a method', *las edificios* 'the buildings', *una camino* 'a path', *una mercado* 'a market'). This parallels the (native) Spanish lexicon: Harris (1991) notes that there is only one non-exotic noun that ends in *-o* but is feminine (*la mano* 'hand'). Harris further reports that the Spanish lexicon contains nearly 600 instances of nouns that end in *-a* but are masculine.

A common error of feature clash involves these exceptional masculine nouns that end in *-a* (*programa* 'program', *sistema* 'system'). These account for six of the 13 errors of feature clash involving feminine determiners in a masculine context. They can most likely be attributed to an overgeneralization of an "*-a* nouns are feminine" strategy. Thus it is possible that some subjects have not yet learned the exceptions to this generalization, and consider the nouns to be feminine, in which case no agreement error is involved (see discussion).

Both the intermediate-level subjects and the advanced-level subjects produced errors in gender, indicating that gender remains a persistent problem (as previously noted by White et al 2004, and Franceschina 2001). Of the 13 errors of feature clash, 3 were produced by advanced speakers; of the 89 errors of underspecification, 23 were produced by advanced speakers. I will further speculate on the relationship between proficiency level and error type in the discussion of the elicited production experiment reported in this chapter. At this point, I will simply point out that morphological variability in gender persists even in the speech of advanced speakers. Individual results for gender in spontaneous production are reported in Appendix D.

5.2.3 Discussion

The data support the hypothesis that errors are ones of underspecification, under the assumption that masculine and singular are underspecified. The pattern of errors supports the claim that features are asymmetrically represented in the lexicon, and that masculine and singular are systematically chosen as defaults due to this asymmetrical representation. If, on the other hand, all of the features under consideration (masculine, feminine, singular, and plural) were represented equally, we would have no reason to expect the errors to be unidirectional.

By studying gender agreement in spontaneous production, we face one important limitation. Implicitly, I have assumed that the errors that have surfaced are, in fact, errors. However, when we examine the data from the perspective of the L2er's grammar—as opposed to the grammar of a native speaker of Spanish—this is assumption is not necessarily warranted. When an error in determiner agreement occurs, it can potentially be either one of two types. The error I describe in (a) is not necessarily an error from the L2er's perspective, whereas the error in (b) is truly an error:

a) the gender of the vocabulary item is incorrectly learned/encoded (for example, the L2er believes that a feminine noun is actually masculine, in which case the L2er has not committed an error from the standpoint of his/her own grammar);

b) gender of the item is correctly learned/encoded, but incorrect morphology is erroneously produced in place of the target form (the L2er knows that a feminine item is feminine, but uses a masculine determiner for whatever reason).

The error in (b) is truly an error because a mismatch would have surfaced between the feminine syntactic context—established by the knowledge that the head N is feminine—and the masculine determiner. When used in isolation, spontaneous production as a methodology makes it impossible to distinguish between these two cases. Ideally, we should look at only (b)-type errors in describing default inflection, since (a)-type errors are more properly an issue of vocabulary learning, and do not tell

us much about the mechanism of agreement. Many—if not most—of the previous studies of the acquisition of gender agreement in Spanish have not attempted to distinguish between these two types of error (e.g. Franceschina 2001, 2002). White et al (2004) recognize the problem and make an attempt at establishing gender knowledge by using a vocabulary test where subjects provided both the name and gender of objects that were targeted in a comprehension task. I follow in adopting a similar vocabulary test, described in the sections that follow.

In principle, there is no reason for incorrect learning/encoding to favor one gender over another[44], and thus it seems unlikely that the asymmetrical relationship between features is due to (a)-type errors alone. Nevertheless, the uncertainty regarding the types of errors remains a problem. The following experiments, outlined in 5.2.2 and 5.2.3, seek to eliminate this problem, or at least control for it as much as possible.

5.3 Elicited production: Gender and number agreement in clitics and adjectives

This section describes an experiment that looks at other domains of agreement: direct object (DO) clitics, and predicative and attributive adjectives (PA, AA). Examples of these domains are given below.

DO clitics, like determiners, are inflected for gender and number agreement with their referent. DO clitics were chosen because I suspected they might be 'harder' for L2ers than determiner agreement and therefore show more errors: the determiner is part of the NP and usually immediately adjacent to the head noun, whereas the DO clitic substitutes for an entire NP. (5) is an example of two sentences: the first contains a full NP (*una manzana*), and the second a DO clitic (*la*).

5. Tiene una manzana. Está comiéndo**la**.
   Have-3SG an-FEM apple. Is eating-CL-FEM
   'He has an apple. He's eating it.'

[44] More properly, there is no explanation within generative theory. Of course, frequency may play a role in the learning or encoding of gender, but this is not something we can model in generative theory.

The use of a DO clitic versus a full NP is constrained by discourse factors, in that noun phrases that are actively in consideration due to recent mention are likely to be realized as clitics. In order to make the DO active in the discourse, a series of two questions was asked. The first question prompted the subject to name the NP, and the second prompted the subject to talk about that NP as the DO of a transitive verb in a sentence. The second question usually came immediately after the first, in order to maintain activation of the DO and create a context for DO clitic use.

Adjectives in Spanish also agree in gender and number with the head noun, which may or may not be overt. Adjectives were further broken down into predicative (a) vs. attributive (b) adjectives, as in (6). This breakdown was done to test whether agreement was more accurate in AAs, as previously reported in the literature (e.g. Bruhn de Garavito & White 2002).

6. a) La camisa es blanca.
   The-FEM shirt is white-FEM
   b) Lleva una camisa blanca.
   She's wearing a-FEM shirt white-FEM

5.3.1 Method

Subjects

Subjects were recruited from Spanish language courses at a major Ontario university. All were native English speakers. They had first been exposed to Spanish in their early teens or later; making them post-critical-period learners. Most subjects reported having had some exposure to French as a subject in school, but none had participated in French immersion programs, and none reported being a bilingual in French and English. Spanish is therefore the L3 of most subjects. The subjects in this experiment also participated in the comprehension experiment described in Section 4.3. The comprehension task was presented first, followed by the production task.

Proficiency was measured according to the same proficiency test described in Section 4.1. Only those subjects that scored in the intermediate and advanced range are included in the sample.

A total of 24 subjects were included in the analysis. 9 scored in the advanced range, and 15 in the intermediate range. Two subjects were excluded due to low proficiency; these subjects also had considerable difficulty completing the elicited production task as they lacked basic vocabulary.

Materials

Test materials consisted of 20 color photographs. These photographs pictured an agent acting upon an object—the latter intended to be realized as a DO clitic. For example, one picture displayed a boy holding an apple up to his mouth, about to bite into it. Another displayed two girls, the first holding a pair of used textbooks, the second holding money, apparently about to buy the books. The objects were chosen with the goal of eliciting 10 masculine noun phrases and 10 feminine ones. In addition, 10 of the objects were intended to elicit singulars, and 10 to elicit plurals. Sample photographs are found in Appendix K.

Table 3. Elicited production task items by gender and number

	Masculine	**Feminine**
Singular	el arete 'the earring' el café 'the coffee' el pájaro 'the bird' el libro 'the book' el periódico 'the newspaper'	la manzana 'the apple' la carta 'the letter' la taza 'the cup' la ventana 'the window' la pelota 'the ball'
Plural	los cuadernos 'the notebooks' los libros 'the books' los zapatos 'the shoes' los pantalones 'the pants' los lápices 'the pencils'	las tijeras 'the scissors' las botas 'the boots' las hojas 'the leaves' las revistas 'the magazines' las camisas 'the shirts'

Procedure

Participants were interviewed individually by a native speaker of Spanish. Each participant was shown all 20 pictures, one at a time. Order of presentation was randomized. The experimenter initially asked a question designed to prompt the participant to name the target noun phrase. Questions for two sample items are shown in (7):

7. a) Qué tiene el chico en la mano?
   What has the boy in the hand?
   b) Qué tiene la chica en la mano?
   What has the girl in the hand?

The participant would typically respond by naming the item, usually along with a determiner inflected for gender and number, as in (8):

8. a) Una manzana.
   An-FEM apple
   b) Unos libros.
   Some-MASC books.

Immediately after naming the object, the experimenter asked a question intended to elicit a clitic, as in (9). This question obligatorily followed the naming of the object in order establish a context for a clitic.

9. a) ¿Qué va a hacer con la manzana?

What going-to do with the-FEM-SG apple?

'What is he going to do with the apple?'

b) ¿Qué está haciendo con los libros?

What is doing with the-MASC-PL books?

'What is she doing with the books?'

The response to the questions in (9) elicited a clitic in some cases, as in (10a). However, in some cases the participant used a full noun phrase, as in (10b):

10. a) Va a comerla.

Going-to eat-CL-FEM-SG

'He is going to eat it.'

b) Está vendiendo los libros.

Is selling the-MASC-PL books

'She is selling the books.'

Following the description of the action in (10), the experimenter asked questions about the color of the object, and of other objects in the photo.

11. a) De qué color es la manzana?

Of what color is the-FEM-SG apple?

'What color is the apple?'

Es roja.

Is red-FEM-SG

'It's red.'

b) De qué color son sus pantalones?

Of what color are her-PL pants?

'What color are her pants?'

Son negros.

Are black-MASC-PL

'They're black.'

This question was designed to elicit adjectives as another domain of gender and number agreement. The interviewer then asked a variety of other questions: *where is the girl?, what else do you see in the picture?, what clothes is she wearing?,* etc. in order to serve as distracters from the test questions. This procedure was repeated for all 20 pictures. Interviews lasted between 10 and 30 minutes.

Data Analysis

All interviews were transcribed by a near-native speaker of Spanish. Productions of clitics and adjectives were coded twice: once for accuracy of gender agreement, and once for accuracy of number agreement. It was also noted whether or not the subject had previously named the target object with the correct gender on the determiner. Two separate analyses are provided below for both clitics and adjectives: one for all of the items, and one for only those that the subject was able to correctly name and provide the gender of. I take 100 percent accuracy in determiner usage to be a diagnostic of knowledge of appropriate gender, and I will refer to this variable as "known gender", as opposed to simply "gender", the sum of all tokens.[45] While determiner usage is in fact variable, and may remain variable even when gender is known (see Section 5.2.3), the exclusion of items where gender is clearly *not* known

[45] When subjects failed to demonstrate knowledge of gender via determiner usage, I was able to establish knowledge of gender by consulting the vocabulary test for some items that appeared on this task. 6 of the 10 feminine items and 8 of the 10 masculine ones appeared on the vocabulary test.

is at least a better approximation to the set of known items. "Known gender" is a considerably smaller data set.

In the analysis of clitics, the subject often failed to provide the expected name of the lexical item. When the subject did not know the name of the target object, the experimenter provided it by giving the name of the noun.[46] These cases are included in the first analysis of all produced clitics, but excluded from the second analysis of only clitics referring to correctly-named objects of "known gender". In other cases, a different name was provided than the intended one. For example, *pelota* 'ball' was often substituted with *bola*. In some cases the target gender remained the same despite the substitution. Where an alternative name and gender were offered, the item was moved to the other category. For example, in the picture showing a girl putting on shoes, some subjects called these *zapatillas*-FEM instead of the intended *zapatos*-MASC. This type of category change led to an unequal number of masculine and feminine test items for some speakers.

Some of the clitic test items also changed number categories. The plural items *zapatos* 'shoes' and *botas* 'boots' were sometimes identified as singular items. For example, for the picture that involved a girl putting on her boots, the participant sometimes said that she was putting on *a boot*, rather than *boots*. The picture showed two boots, but the girl was holding only one shoe in the photograph, and so these items were counted as appropriate use of singular inflection where the clitic was singular.

In the analysis of adjectives, those that are homophonous for gender (e.g. *verde* 'green', *azul* 'blue') were excluded from gender analysis, but included for number analysis where they do bear overt number agreement (*verdes* 'green-PL', *azules* 'blue-PL'). Some adjectives bear agreement for neither gender nor number (e.g. *beige* 'beige', *naranja* 'orange', *rosa* 'pink'), and these were completely excluded from analysis. Conjoined adjectives were counted once when agreement was consistently

[46] The experimenter only supplied the word when it was clear that the subject did not know it, which was indicated in a variety of ways (e.g. asking for the name, or saying it in English). In some cases, the experimenter happened to provide the name of the object alone (*tijeras* 'scissors') and in other cases with a determiner (*las tijeras* 'the scissors'). Ideally, the experimenter's degree of helpfulness could have been more tightly controlled; however, this issue was not anticipated prior to data collection.

targetlike or nontargetlike for each adjective (e.g. *blanco y negro* 'white-MASC' and black-MASC'). Where the conjoined adjectives differed in agreement properties (e.g. *blanca y negro* 'white-FEM and black-MASC'), each adjective was coded separately. Adjectives that were part of null nominal phrases (*la otra* 'the other one') were excluded, as the absence of a noun presented an additional confounding factor (though see White et al 2004 for an analysis of gender and number in null nominals).

Statistical analysis

For both clitics and adjectives, the percent accuracy was tallied for each feature (masculine, feminine, singular, plural) and each category (gender, number). A series of chi-square tests were performed to test for the independence of independent and dependent variables (masculine, feminine gender vs. accurate, inaccurate; singular, plural number vs. accurate, inaccurate; intermediate, advanced, NS group vs. accurate, inaccurate). The number of tokens elicited for each participant varies widely; some subjects used very few clitics and/or adjectives, whereas some used them consistently. For this reason, data is analyzed as counts rather than rates. In addition, the effect of knowledge of gender is calculated for each category and gender feature.

5.3.2 Results

5.3.2.1 Clitics

Accuracy and group

The elicited production task elicited a total of 390 tokens of gender agreement, and 405 tokens of number agreement in clitics from the L2 subjects. Table 4 summarizes the accuracy rates for gender agreement in clitics by proficiency group and feature. Individual results are reported in Appendix E.

Table 4. Percent accuracy in clitic agreement, by group and feature/category

	All L2	Advanced	Intermediate	Natives
Gender (All)	83.3	86.9	80.9	100.0
Masculine	91.7	92.3	91.3	100.0
Feminine	75.1	81.7	70.4	100.0
Number (All)	98.3	98.8	98.0	100.0
Singular	100.0	100.0	100.0	100.0
Plural	96.3	97.4	95.6	100.0

Overall, subjects are more accurate with number than with gender agreement (98 percent vs. 83 percent), as expected based on previous research. The L2 subjects show 100 percent accuracy for singular agreement, and high rates of accuracy for plural agreement as well (about 96 percent).

The contingency between group (NS, advanced, intermediate) and accuracy is significant for gender ($X^2 = 43.51$, df = 1, $p<.0001$). Gender accuracy increases with proficiency level, and NSs are more accurate than both L2 groups. Gender and number accuracy by group are shown in Tables 5 and 6. There are too few tokens of number errors to warrant testing via chi-square analysis, as expected cell values are below 5.[47] The difference between the NS group and both the advanced and intermediate groups is significant for gender ($X^2 = 52.31$, df = 1, $p <.0001$ for NS vs. advanced; $X^2 = 96.67$, df = 1, $p < .0001$ for NS vs. intermediate), as is the difference between advanced and intermediate ($X^2 = 44.76$, df = 1, $p < .0001$).

[47] Various sources attest to the invalidity of chi-square tests when expected cell totals are small, including Freund and Simon (1995, p. 411): "Since the sampling distribution of the X^2 statistic we are using here is only approximately a chi-square distribution, it should not be used when any of the expected cell frequencies are less than 5."

Table 5. Gender accuracy versus group

	Intermediate	Advanced	Native	Totals
Accurate	186	139	214	539
Inaccurate	44	21	0	65
Totals	230	160	214	604

Table 6. Number accuracy versus group

	Intermediate	Advanced	Native	Totals
Accurate	240	158	219	617
Inaccurate	5	2	0	7
Totals	245	160	219	624

Clitics: Gender Agreement by Feature

A total of 193 clitics were used in referring to masculine NPs, and 197 were used in referring to feminine NPs. Tables 7, 8, and 9 are contingency tables that show the distribution of gender value by accuracy.

Table 7. Gender agreement in clitics for all L2 subjects: feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	177	148	325
Inaccurate	16	49	65
Totals	193	197	390

Table 8. Gender agreement in clitics for advanced L2 subjects: feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	72	67	139
Inaccurate	6	15	21
Totals	78	82	160

Table 9. Gender agreement in clitics for intermediate L2 subjects: feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	105	81	186
Inaccurate	10	34	44
Totals	115	115	230

Subjects produce more errors in feminine contexts. The contingency between feature value and accuracy is significant for the L2 subjects overall (X^2= 36.22; df = 1, p<.0001). The contingency is not significant for the advanced group (X^2= 2.50, df = 1, p < .2). The advanced subjects were quite accurate with gender, producing a total of 21 gender errors out of 160 clitics; thus the sample is quite small and the failure to reach significance is likely an effect of the low error rate. The intermediate group, on the other hand, produced 44 errors out of 230 tokens, and show a contingency between feature value and accuracy, with more errors in feminine contexts (X^2 =

10.26, df = 1, p<.01). The effect of gender for the entire L2 sample can be attributed largely to the behavior of the intermediate subjects.

Knowledge

At this point, I will limit the data set to items of known gender; that is, I will count only those NPs the subject was able to name along with the appropriate gender. This will come closer to separating errors classified as 'true' errors (those items for which the subject demonstrates accurate gender knowledge, but uses a clitic with nontargetlike gender: type-b errors) from other errors (those items which the subject potentially has miscoded for gender: type-a errors).

In addition to the effect of feature, there appears to be a consistent effect of knowledge of gender on accuracy, as accuracy rates climb when only known gender is considered. Masculine agreement hovers around 92 percent for all groups for overall gender, but this figure jumps to 100 percent for the advanced group, and 93.1 percent for the intermediate group, for known gender tokens. Thus, for advanced speakers, the errors in masculine gender surface only in cases where the subject has either not demonstrated knowledge, or has demonstrated incorrect knowledge, of the gender of the test item. Accuracy rates for known gender are shown in Table 10.

Table 10. Percent accuracy in clitic agreement for known gender by group and feature/category

	All L2	Advanced L2	Intermediate L2
Gender (All)	85.9	90.6	82.5
Masculine	95.7	100.0	93.1
Feminine	75.7	82.5	69.9

For known gender, a significant contingency between gender and accuracy is found for the L2 subjects as a whole ($X^2 = 14.38$, df = 1, $p<.0002$), and for each proficiency group considered separately (intermediate: $X^2 = 10.69$, df = 1, $p<.002$; advanced: $X^2 = 12.08$, df = 1, $p<.001$). Recall that no significant contingency was found for the advanced group when all gender items were considered; by limiting the data set to known gender, the effect of feature emerges. The advanced group produced only 11 errors in the category of known gender, but all involved the substitution of a masculine clitic in a feminine context. These results indicate that under both analyses (total production of gender in clitics, and known gender in clitics), subjects are significantly more accurate with clitics in masculine contexts than feminine contexts. Tables 11, 12, and 13 show known gender by accuracy.

Table 11. Known gender agreement in clitics for all L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	135	103	238
Inaccurate	6	33	39
Totals	141	136	277

Table 12. Known gender agreement in clitics for advanced L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	54	52	106
Inaccurate	0	11	11
Totals	54	63	117

Table 13. Known gender agreement in clitics for intermediate L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	81	51	132
Inaccurate	6	22	28
Totals	87	73	160

A chi-square analysis of knowledge of the gender of masculine and feminine items reveals a significant contingency between accuracy and knowledge ($X^2 = 18.35$, df = 1, $p < .0001$). This contingency is shown in Table 14. For masculine gender alone, there is a significant contingency, such that subjects are more accurate with masculine gender when they know the object is masculine ($X^2 = 5.72$, df = 1, $p<.02$); see Table 15. Put differently, when subjects know an object to be masculine, they generally do not use a feminine clitic.

Feminine agreement, unlike masculine agreement, remains variable when data are limited to items where gender is known (75.7 percent overall; 82.5 for advanced; 69.9 for intermediate). Unlike masculine agreement, there is no significant contingency between knowledge of feminine gender and accuracy ($X^2 = 2.17$, df = 1, $p<.2$). The contingency between knowledge of feminine gender and accuracy is shown in Table 16. The contrast in accuracy rates when gender is known suggests a fundamental asymmetry between the two features in opposition: errors in masculine agreement surface when the subject does not know the gender of the target item, whereas errors in feminine agreement surface both when gender is known and when it is not known.

Table 14. Knowledge of gender by accuracy in clitics.

	Knowledge	No knowledge	Totals
Accurate	238	87	325
Inaccurate	39	26	65
Totals	277	113	390

Table 15. Knowledge of masculine gender by accuracy in clitics.

	Knowledge	No Knowledge	Totals
Accurate	135	42	177
Inaccurate	6	10	16
Totals	141	52	193

Table 16. Knowledge of feminine gender by accuracy in clitics.

	Knowledge	No Knowledge	Totals
Accurate	103	45	148
Inaccurate	33	16	49
Totals	136	61	197

Table 17 shows the distribution of errors in a different way. Here, the data are limited to errors alone, and broken down by gender category and knowledge. Out of the 39 errors of known gender that emerged in the data set, only 6 involved a

masculine context. The majority of errors of known gender involve the use of a masculine clitic in a feminine context. The contingency between knowledge of gender and gender feature is significant ($X^2 = 26.82$, df = 1, $p<.0001$). I therefore conclude that the majority of "true" errors (i.e. type-b: those errors that involve nontargetlike agreement despite the correct knowledge of a noun's gender) involve the use of masculine clitics in feminine contexts.

Table 17. Errors by gender and knowledge (all L2 subjects)

	M	F	Totals
No knowledge	10	16	26
Knowledge	6	33	39
Totals	16	49	65

When we break these errors down even further, we see that it is the intermediate L2ers who commit these 6 errors of masculine known gender. Recall that the advanced L2ers' accuracy rate in known gender was 100 percent (see Table 13). Advanced L2ers still make errors, but they all involve feminine known gender, whereas intermediate L2ers make errors involving both masculine and feminine known gender. I will speculate as to what this effect derives from in Chapter Six.

Number

The elicited production task elicited a total of 405 number agreement tokens.[48] Only 7 of these were errors. The data are presented in Tables 18, 19, and 20. Due to the low expected values in each cell, chi-square tests are not performed (see note 47).

[48] This figure is not identical to the number of gender tokens largely because of the subjects' tendency to reduce vowels. Productions of phonetic [lə] or [ləs] were coded for number, but not for gender.

Table 18. Number agreement in clitics for all L2 subjects: Feature versus accuracy.

	Singular	Plural	Totals
Accurate	215	183	398
Inaccurate	0	7	7
Totals	215	190	405

Table 19. Number agreement in clitics for advanced L2 subjects: feature versus accuracy.

	Singular	Plural	Totals
Accurate	83	75	158
Inaccurate	0	2	2
Totals	83	77	160

Table 20. Number agreement in clitics for intermediate L2 subjects: feature versus accuracy.

	Singular	Plural	Totals
Accurate	132	108	240
Inaccurate	0	5	5
Totals	132	113	245

Although the error rate is extremely low, the data show a complete absence of plural morphology in singular contexts. This is true of both proficiency groups.

5.3.2.2 Adjectives

Accuracy and group

The elicited production task elicited a total of 378 tokens of gender agreement, and 488 tokens of number agreement in adjectives from the L2 subjects. Table 21 summarizes the accuracy rates for gender and number agreement in adjectives by proficiency group and feature. Individual results are reported in Appendix F.

Table 21. Percent accuracy in adjective agreement, by group and feature/category

	All L2	Advanced	Intermediate	Natives
Gender (All)	79.1	88.6	73.5	100
Masculine	92.8	94.2	91.9	100
Feminine	66.7	83.1	57.5	100
Number (All)	93.9	96.8	92.1	99.4
Singular	99.7	100.0	99.6	100.0
Plural	73.4	87.5	62.3	97.3

Subjects are less accurate with gender agreement than with number agreement (79.1 percent for gender vs. 93.9 percent for number). This parallels the pattern reported for clitics in Table 1.

There is a significant contingency between group and accuracy for both gender ($X^2 = 24.76$, df = 2, $p < .0001$) and number ($X^2 = 7.98$, df = 2, $p<.005$). Natives are more accurate than the advanced group, who are more accurate than the intermediate group.

All group comparisons are significant for gender (NS vs. advanced: $X^2 = 26.20$, df = 1, $p < .0001$; NS vs. intermediate: $X^2 = 88.98$, df = 1, $p < .0001$; advanced vs. intermediate: $X^2 = 62.14$, df = 1, $p < .0001$) and number (NS vs. advanced: $X^2 = 9.13$, df = 1, $p < .003$; NS vs. intermediate: $X^2 = 36.27$, df = 1, $p < .0001$; advanced vs. intermediate: $X^2 = 33.35$, df = 1, $p < .0001$). Contingency tables for gender and number accuracy across groups are shown in Tables 22 and 23.

Table 22. Gender accuracy in adjectives versus group

	Intermediate	Advanced	Natives	Totals
Accurate	175	124	140	439
Inaccurate	63	16	0	79
Totals	238	140	140	518

Table 23. Number accuracy in adjectives versus group

	Intermediate	Advanced	Natives	Totals
Accurate	279	179	173	631
Inaccurate	24	6	1	31
Totals	303	185	174	662

Adjectives: Gender Agreement by Feature

Of the 378 tokens of gender agreement in adjectives, 79 were errors. Contingency tables for gender features by accuracy are given in Tables 24, 25, and 26.

Table 24. Gender agreement in adjectives for all L2 subjects: feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	167	132	299
Inaccurate	13	66	79
Totals	180	198	378

Table 25. Gender agreement in adjectives for advanced L2 subjects: feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	65	59	124
Inaccurate	4	12	16
Totals	69	71	140

Table 26. Gender agreement in adjectives for intermediate L2 subjects: feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	102	73	175
Inaccurate	9	54	63
Totals	111	127	238

The contingency between gender and accuracy is significant for the L2 subjects overall ($X^2 = 48.76$, df = 1, $p<.0001$) and for the intermediate subjects ($X^2 = 24.60$, df = 1, $p<.001$), with more errors occurring in feminine contexts. The advanced L2 group's contingency does not reach significance though it is indicative of a trend ($X^2= 3.62$, df = 1, $p<.06$).

For adjectives agreeing with nouns of known gender, a total of 46 gender errors out of 273 tokens were recorded. Again, we find a significant effect overall ($X^2 = 20.33$, df = 1, $p<.0001$; see Table 27) and for intermediate subjects ($X^2 = 20.70$, df = 1, $p<.0001$; see Table 29). Only 10 errors were found for the advanced subjects (Table 28), too few to analyze using a chi-square statistic (see note 4); nevertheless the data suggest an asymmetry in known gender between masculine and feminine contexts, as 8 of 10 errors occur in feminine contexts.

Table 27. Known gender agreement in adjectives for all L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	115	112	227
Inaccurate	4	42	46
Totals	119	154	273

Table 28. Known gender agreement in adjectives for advanced L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	44	53	97
Inaccurate	2	8	10
Totals	46	61	107

Table 29. Known gender agreement in adjectives for intermediate L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	71	59	130
Inaccurate	2	34	36
Totals	73	93	166

Adjectives: Attributive vs. Predicative

Adjectives can be further broken down into types: attributive adjectives (AAs) and predicative adjectives (PAs; see example 6). Only L2 data are reported in this section, as NS accuracy was 100 percent across all adjectives (except for one error in plural number agreement—a PA). PAs yielded 281 tokens of gender agreement and 373 tokens of number agreement from L2 speakers. AAs yielded 97 tokens of gender agreement and 115 tokens of number agreement.

AA agreement is more accurate than PA agreement, in accordance with other studies on L2 adjective agreement (e.g. Bruhn de Garavito & White 2002). Here, AAs are more accurate than PAs for both gender (91 vs. 75 percent for gender overall) and

number (98 vs. 93 percent). Table 31 presents the distribution of errors for AAs—since only 9 errors surfaced, attributive errors do not provide a rich source of data and will not be discussed further. Contingency tables for features in predicative adjectives are given in Tables 32 and 33.

Table 30. Percent accuracy in attributive (A) and predicative (P) adjective agreement, by group and feature/category

	All L2		Advanced		Intermediate	
	A	**P**	**A**	**P**	**A**	**P**
Gender	90.7	75.0	95.0	86.0	87.7	69.1
Masculine	92.7	92.8	92.9	94.5	92.6	91.7
Feminine	89.3	57.7	96.2	75.6	83.3	49.5
Number	98.3	92.5	97.8	96.4	98.6	90.1
Singular	100.0	99.7	100.0	100.0	100.0	99.5
Plural	94.1	64.0	92.9	85.3	95.0	46.3

Table 31. Gender agreement in attributive adjectives for all L2 subjects: Feature versus accuracy.

	Masculine	**Feminine**	**Totals**
Accurate	38	50	88
Inaccurate	3	6	9
Totals	41	56	97

Table 32. Gender agreement in predicative adjectives for all L2 subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	129	82	211
Inaccurate	10	60	70
Totals	139	142	281

Table 33. Gender agreement in predicative adjectives for all advanced subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	52	34	86
Inaccurate	3	11	14
Totals	55	45	100

Table 34. Gender agreement in predicative adjectives for intermediate subjects: Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	77	48	125
Inaccurate	7	49	56
Totals	84	97	181

The contingency between gender and accuracy is upheld for PAs ($X^2 = 32.56$, df = 1, p<.0001; see Table 32). Both groups show a contingency between gender feature and accuracy for PAs when considered separately ($X^2 = 11.72$, df = 1, p<.001 for advanced, Table 33; $X^2 = 26.72$, df = 1, p<.0001 for intermediate, Table 34). The contingency between gender and accuracy appears to extend to advanced subjects when PAs are singled out for analysis. Recall that no such contingency was found for the advanced subjects when both types of adjectives were considered together. Thus it is clear that in production, variability in gender agreement, in particular the use of default inflection, persists even for speakers at high levels of proficiency in this domain of the grammar.

Knowledge

Percent accuracy in adjective agreement with nouns of known gender across groups is shown in Table 35. Restricting the data set to known gender items has a positive effect on accuracy rates, though variability remains. The main effect of knowledge of gender on accuracy is significant ($X^2 = 19.05$, df = 1, p<.0001; see Table 36), with gender accuracy higher for items where subjects have demonstrated knowledge of noun gender. Subjects were highly accurate with masculine known gender, with only 13 errors out of 180 masculine contexts. The effect of knowledge of gender on masculine items is significant ($X^2 = 4.40$, df = 1, p<.04); see Table 37. The contingency between feminine gender and accuracy is also significant for adjectives ($X^2 = 5.64$, df = 1, p<.02); see Table 38.

Table 35. Percent accuracy in adjective agreement for known gender by group and feature/category

	All L2	Advanced L2	Intermediate L2
Gender (All)	83.2	90.7	78.3
Masculine	96.6	95.7	97.3
Feminine	72.7	86.9	63.4

Table 36. Knowledge of all gender by accuracy in adjectives.

	Knowledge	No knowledge	Totals
Accurate	227	72	299
Inaccurate	46	33	79
Totals	273	105	378

Table 37. Knowledge of masculine gender by accuracy in adjectives.

	Knowledge	No knowledge	Totals
Accurate	115	52	167
Inaccurate	4	9	13
Totals	119	61	180

Table 38. Knowledge of feminine gender by accuracy in adjectives.

	Knowledge	No knowledge	Totals
Accurate	112	20	132
Inaccurate	42	24	66
Totals	154	44	198

By isolating the errors and breaking them down by knowledge and gender feature, the effect of gender clearly emerges. Out of the 46 errors involving items of known gender, only 4 of these occurred in a masculine contexts. The contingency between knowledge of gender and gender feature is significant ($X^2 = 19.38$, df = 1, $p<.0001$; see Table 39). As was found for clitics, the use of feminine clitics in masculine contexts is rare, and is especially rare when the subject knows that the target item is masculine.

Table 39. Errors in adjectives by gender and knowledge (all L2 subjects)

	Masculine	Feminine	Totals
No knowledge	9	24	33
Knowledge	4	42	46
Totals	13	66	79

For adjectives, we do not see the same relationship between proficiency level and errors of known gender that we saw for clitics. It appears that both the intermediate and advanced groups are highly accurate with masculine known gender, though the advanced group does not reach 100 percent accuracy in masculine known gender, as we saw for clitics.

Number

A total of 488 tokens of number agreement in adjectives were recorded. 30 of these were errors, and 29 of these involved the substitution of a singular adjective for a plural one. A significant contingency is obtained between number feature and accuracy overall (X^2 = 28.19, df = 1, p<.0001; see Table 40), and within the intermediate group (X^2 = 36.81, p<.0001 for intermediate; see Table 42). Again, the error rate for advanced subjects yields a data set that is too small to test for a feature-accuracy contingency; nevertheless, the 6 errors that occur in number agreement are all found in plural contexts (see Table 41). 28 of the 30 errors are found in predicative adjectives.

Table 40. Number agreement in adjectives for all L2 subjects:

Feature by accuracy

	Singular	Plural	Totals
Accurate	378	80	458
Inaccurate	1	29	30
Totals	379	109	488

Table 41. Number agreement in adjectives for advanced subjects:

Feature by accuracy

	Singular	Plural	Totals
Accurate	137	42	179
Inaccurate	0	6	6
Totals	137	48	185

Table 42. Number agreement in adjectives for intermediate subjects: Feature versus accuracy.

	Singular	Plural	Totals
Accurate	241	38	279
Inaccurate	1	23	24
Totals	242	61	303

5.3.3 Discussion

The results of the elicited production experiment detailed in this section parallel the findings of the spontaneous production study presented in Section 4.1. In both of these experiments, gender was found to be less accurate than number, and masculine inflection was used systematically as a default. In light of the data from these two studies of the production of gender and number agreement, three major observations can be made: morphological variability is persistent, it is systematic, and it is affected by lexical knowledge of gender.

First, the results of the spontaneous and elicited production studies confirm that variability for many speakers is a persistent phenomenon found at even very high levels of proficiency. This observation adds to the body of research on persistent L2 morphological variability, including Lardiere (1998) and White (2003) for English, and Franceschina (2001) for Spanish. PA agreement turned out to be highly problematic even for advanced learners, who correctly used feminine gender in feminine contexts at a rate of only 75.6 percent, and plural agreement in plural contexts at a rate of 85.3 percent. Of the nine advanced speakers who participated in the elicited production task, only two showed 100 percent accuracy in agreement across the board.

The following are examples of gender errors in clitics (12), and adjectives (13, 14) made by speakers at advanced levels of proficiency. The examples include dialogue

between the experimenter (E) and the subject (S) that demonstrates the subject's knowledge of the gender of the target item.

Subject 27, Advanced L2 Spanish

12. E: Qué está tocando el chico con las manos? ("What is the boy touching with his hands?")

    S: **La ventana**. ("The-FEM window.")

    E: Qué es lo que quiere hacer con la ventana? ("What is it that he wants to do with the window?")

    S: **Lo** quiere cerrar.

    CL-MASC.SG wants to-close

    "He wants to close it."

Subject 16, Advanced L2 Spanish

13. S: Tiene tijeras, y está poniendo las tijeras en **la mochila**. [...] ("She has scissors, and she's putting the scissors in the-FEM backpack.")

    E: De qué color es la mochila? ("What color is the backpack?")

    S: **La** **mochila** es **negro** con un poco de blanco.

    The-FEM backpack is black-MASC with a little white

    "The backpack is black with a little white."

Subject 33, Advanced L2 Spanish

14. S: Está poniendo **las zapatillas.** ("She's putting the-FEM-PL sneakers.")

    E: De qué color son las zapatillas? ("What color are the sneakers?")

    S: **Blancos.**

    White-MASC-PL

Gender errors in (12-14) contrast with the appropriate use of gender by the same subjects in (15-17). Together, these examples demonstrate morphological variability.

Subject 27, Advanced L2 Spanish:

15. E: Este chico, qué tiene en la mano? ("This boy, what does he have in his hand?")

    S: **Una** manzana.

    An-FEM apple

    E: Qué está haciendo con la manzana? ("What is he doing with the apple?")

    S: **La** va a comer.

    CL-FEM going-to eat

    "He's going to eat it."

Subject 16, Advanced L2 Spanish:

16. E: De qué color es la mesa? ('What color is the table?')

    S: La mesa es blanca.

    The-FEM table is white-FEM

    "The table is white."

Subject 33, Advanced L2 Spanish:

17. E: Cómo es la ropa de las chicas? ('What is the girls' clothes like?')

    S: Tiene pantalones blancos, camiseta negra.

    They-have pants-MASC white-MASC-PL, t-shirt-FEM black-FEM

Number agreement also remains somewhat variable for advanced speakers, particularly in the domain of predicative adjectives. In (18), the subject uses singular agreement in a plural context. Here, the use of a plural copula *son* makes the syntactic context unambiguously plural. At the same time, the subject uses masculine gender in a feminine context. (Gender of this particular item was not demonstrated, thus this is not an error of 'known gender'.)

Subject 16, Advanced L2 Spanish

18. E: Y te gustan sus medias, o vas a criticarlas? ("And do you like her socks, or are you going to criticize them?")

    S: No no no. Son **gris** y **blanco.**

    No no no. They.are gray-SG and white-SG

    'No, no, no. They're gray and white.'

Subject 16's production of number inflection is truly variable, as she produces correct plural agreement just a few lines of dialogue later:

19. E: Entonces, de qué color son los zapatos? ("So, what color are her shoes?")

    S: Los zapatos son **blancos** y **negros** creo.

    The-PL shoes are white-PL and black-PL believe-1SG

    'The shoes are white and black I think.'

Second, in addition to its persistence, variability is systematic. Errors are not random, but systematic across a range of nominal domains. In the spontaneous production of determiners, and the elicited production of DO clitics and adjectives, masculine inflection occurs in feminine contexts, whereas the reverse pattern is rarely attested. Across the board, masculine determiners, clitics, and adjectives act as defaults in feminine syntactic contexts; singular determiners, clitics, and adjectives act as defaults in plural contexts. In some domains (e.g. number agreement in attributive adjectives) no evidence of variability is found, making it systematically targetlike. Furthermore, the features that were assumed to be underspecified according to markedness values (see Chapter Two) correspond to the systematically-used defaults. These findings, combined with the major finding of Chapter Three—that underspecified inflection is produced in the verbal domain—support an underspecification-based theory of the second language grammar.

Third, exhibited knowledge of target gender plays a role in determining how much variability surfaces, and in what context it surfaces. For clitics, masculine agreement surfaced as a default in feminine contexts both when gender was known and when it was not. Feminine agreement was essentially limited to cases where the target was known to be feminine—feminine clitics are not used as defaults when the subject knows the target to be masculine. In Section 5.2.3 I noted that, in principle, two types of errors were possible: "true" errors in which the subject knows the gender of a noun but gets the agreement wrong, and "false" errors in which the subject believes the item to be of the opposite gender. By breaking down errors according to whether or not subjects demonstrated knowledge of gender, it is clear that errors in feminine contexts occur regardless of knowledge, but errors in masculine contexts occur primarily when the subject does not know the gender of the target item. This pattern suggests that feminine clitics are not used as true defaults, but rather surface when the subject has incorrectly encoded the gender of the target item, or is making a guess at its gender. We might therefore take this pattern to mean that many of the apparent counterexamples to the MUSH actually fall outside of its scope: recall that the MUSH makes predictions about *errors*. If the L2er has incorrectly encoded a given noun as

feminine when it is actually masculine, then the use of feminine agreement with such a noun does not, from the L2er's perspective, constitute an error.

As far as clitics are concerned, this experiment was generally successful in eliciting the target structure. However, there were some subjects that avoided using clitics, and instead chose to repeat the full NP. Below is a passage of an advanced-level subject who avoided clitics during the first half of the experiment:

Subject 16, Advanced L2 Spanish:

20. E: Bueno aquí en esta foto qué ves?

S: Ah, un gatito, o gatita tal vez no sé, y un pájaro en un árbol.

E: Y probablemente qué quiere hacer el gato con el pájaro?

S: Supongo que el gato está mirando el pájaro, pero tal vez quiere comer el pájaro.

E: Sí. ¿Y dónde está el pájaro?

S: Ehm, está en el árbol.

E: ¿De qué color son las hojas del árbol?

S: Parecen verdes.

E: Entonces suponemos que ¿qué va a hacer después el gato con el pájaro?

S: Que tal vez va a saltar, por comer el pájaro.

E: Well here in this picture, what do you see?

S: Ah, a kitty (masc), or kitty (fem) maybe, I don't know, and a bird in a tree.

E: And probably what does the cat want to do with the bird?

S: I suppose that the cat is looking at the bird, but maybe he wants to eat the bird.

E: Yes. And where is the bird?

S: Um, it's in the tree.

E: What color are the leaves on the tree?

S: They look green.

E: Then we suppose, what is the cat going to do with the bird later?

S: That maybe he will jump, to eat the bird.

Over more than one attempt on the part of the experimenter to elicit a clitic, the subject manages to avoid using the structure. Within one utterance, she even repeats the full NP twice in DO position (*el gato está mirando el pájaro, pero tal vez quiere comer el pájaro/the cat is looking at the bird, but maybe he wants to eat the bird).* This creates a somewhat awkward dialogue from the point of view of discourse factors, although each utterance is a grammatical one when discourse factors are not considered. Nevertheless, this passage is unlike the NS passages, where clitics are used consistently. Thus discourse is clearly one domain that remains problematic in that it may resist native-like performance for some L2ers, even at very advanced levels (for a discussion of discourse and L2 Spanish clitics, see for example Torres 2003). At the expense of using discourse-appropriate language, L2ers seem to be avoiding the use of clitics (as they used them very infrequently in spontaneous production), perhaps because of the difficulty of marking them for gender and number agreement.

Another issue that arises is the difference in accuracy between the two types of adjective: PA and AA. Subjects were more accurate with agreement in AA contexts than PA contexts. A likely explanation is the effect of "distance": AAs are closer to the head noun than PAs, which follow a copular verb. If distance completely explains this effect, we might expect for adjective agreement in resultative constructions to be highly accurate (as in *pintar la pared blanca* 'paint the wall red-fem').[49] On the other

[49] I am indebted to Andrea Gualmini for suggesting this point to me.

hand, if L2ers make agreement errors in such constructions, then we might be forced to seek another explanation. This is one potential area for future research.

In this section, I presented the major findings for the production components of this study of gender and number agreement. In the following section, I present the results of a comprehension study on DO clitics in order to assess whether default morphology is limited to production, and whether the same defaults emerge in comprehension as in production. The results of this study will allow us to see to what extent the comprehension and production of L2 agreement are qualitatively different.

5.4 Comprehension: A picture-selection task

This section describes an experiment that examines the comprehension of DO clitics. The goal of this experiment is to complement the spontaneous and elicited production experiments previously discussed in this chapter. In this experiment, the overt expression of gender and number agreement on a DO clitic is the crucial clue to the correct interpretation of a sentence. If default underspecified inflection is found, as predicted, subjects will interpret masculine and singular clitics as appropriate in feminine and plural contexts, respectively.

Subjects that participated in this experiment were the same ones that participated in the elicited production experiment (N=24).

5.4.1 Method

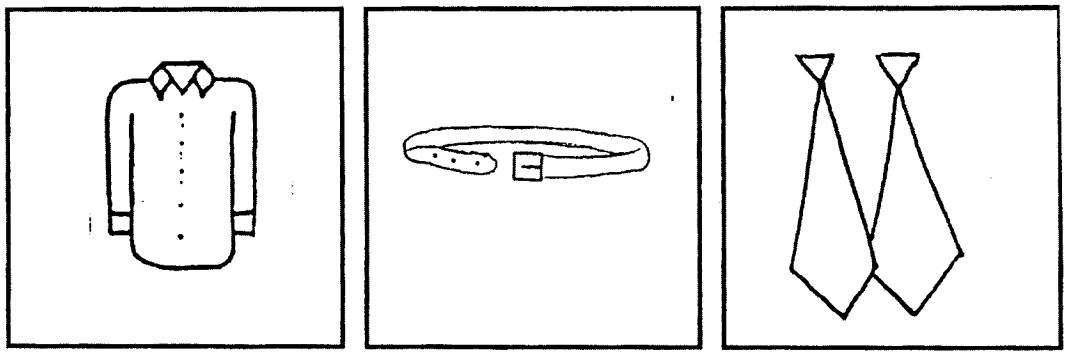
Materials

The materials, adapted from White et al 2004, consisted of a booklet containing a story about two characters going on vacation. The story contained 48 test sentences. 16 of these tested gender (8 masculine clitics, 8 feminine clitics) and 16 of these tested number (8 singular clitics, 8 plural clitics). The remaining 16 were distractors that did not contain clitics (Sample distractor: "Después de un rato, Marta dice: "Paco, vamos a bailar."/*After a while, Marta says: "Paco, let's go dancing."* with pictures of dancing/hiking/singing). After subjects read a sentence containing a clitic,

the subject circled the picture, out of three possible choices, that correctly corresponded to the story. One picture corresponded in gender and number to the clitic. Another picture disagreed in the category of the test item, holding the other category constant. The third picture was a foil: half of the foils disagreed in one category, and half disagreed in two categories. Each picture was designed to be equally plausible in the context of the story. (21) is a sample gender item, and (22) is a sample number item:[50]

21. Paco quiere llevar algunas cosas que acaba de comprar pero no encuentra nada. Paco dice: "Acabo de comprarlo—¿dónde está?"

    Paco wants to bring some things that he just bought, but he can't find anything. Paco says, "I just bought it-MASC—Where is it?"

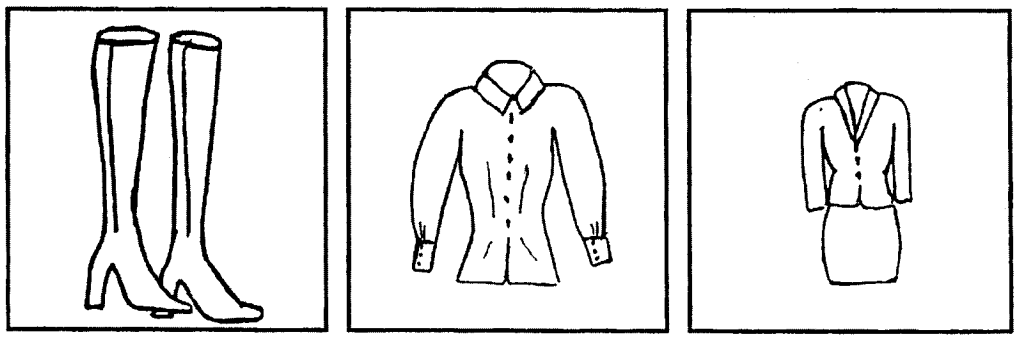


1	2	3
la camisa	el cinturón	las corbatas
the shirt-FEM	the belt-MASC	the ties-MASC-PL

[50] The names of the objects did not appear on the task.

22. Marta dice: "Las voy a llevar a la fiesta de Navidad."

Marta says: "I'm going to wear them-FEM to the Christmas party."



1	2	3
las botas	la blusa	el traje
the boots-FEM-PL	the blouse-FEM	the suit-MASC

The clitic in (21) is masculine and singular: it agrees with *el cinturón*, the target answer. If the subject chose *la camisa* ('the-FEM shirt'), this would constitute an error of underspecification, as the subject selected a feminine context to correspond with a masculine clitic. *Las corbatas* ('the-FEM ties') is the foil—it disagrees in both gender and number. The clitic in (22) is feminine and plural: it agrees with *las botas* ('the-FEM boots'), the target answer. The selection of *la blusa* ('the-FEM blouse') would constitute an error of feature clash, as the subject chose a singular context to correspond with a plural clitic. *El traje* ('the-MASC suit') is the foil for this item, in this case differing in both gender and number. Neither the name nor the gender of the items in the pictures appeared in the story. This ensured that subjects relied on their knowledge of gender and number agreement in order to select the correct item. Knowledge of the name and gender of the test items was independently established by the vocabulary test, described below. See Appendix J for a list of vocabulary.

Subjects were tested individually. The instructions told them that they were going to read a story about two characters going on vacation, and that they should choose

the picture that best corresponded to the story they were reading. The comprehension test was administered after the vocabulary test.

Vocabulary test

The vocabulary test consisted of 48 pictures, all of which appeared in the picture selection task (some as fillers). For each item, subjects wrote the name of the object and circled the article that corresponded in gender. The vocabulary test was administered before the comprehension task.

In some cases, alternate names were given for test items (e.g. *la bola* for *la pelota*—both mean ball though *pelota* is more common); these were included as correct responses. In other cases, alternate names and genders were given for test items (e.g. *la taza* 'the cup' for *el vaso* 'the glass', *el bote* 'the boat' for *la bota* 'the boot'); these items were discarded from the analysis of the comprehension task results, as were items that were left blank.

Data Analysis

On the comprehension task, responses were coded as correct or incorrect in conjunction with responses on the vocabulary test. A correct answer was tallied when the subject knew the gender of the target item (as shown by the vocabulary task) and circled that item. An incorrect answer was tallied when the subject knew the gender of the target item, but chose another item of opposite gender. Answers were excluded from analysis when subject was not able to correctly identify the gender of a given item in the vocabulary test. This was done to attempt to ensure that the subject had a particular name/gender in mind while they were answering the comprehension task.

The number of correct and incorrect responses was tallied for each gender and number feature. A series of chi-square tests were performed to test for contingencies between feature and accuracy (masculine, feminine and singular, plural vs. accurate, inaccurate), and group and accuracy (advanced, intermediate, NS vs. accurate, inaccurate).

5.4.2 Results

Accuracy

The accuracy rates for category and feature by group are presented in Table 43. Intermediate L2 subjects are less accurate than advanced L2 subjects, who are less accurate than the NS group. The intermediate group is significantly less accurate with gender than both of the two other groups ($X^2 = 26.29$, df = 1, $p<.0001$ vs. Advanced; $X^2 = 37.86$, df = 1, $p<.0001$ vs. NS). The difference between the advanced and NS group fails to reach significance ($X^2 = 3.18$, df = 1, $p<.08$). Accuracy rates for gender and number are broken down by group in Tables 44 and 45. Due to the high accuracy rates for number, the distribution of errors by group and accuracy is not large enough to warrant statistical testing. Individual results are reported in Appendix G.

Table 43. Accuracy rates by category/feature and group, comprehension task

	All L2	Advanced L2	Intermed. L2	Natives
Gender (All)	88.9	97.3	82.0	99.3
Masculine	83.2	96.6	72.2	98.6
Feminine	95.0	98.1	92.5	100.0
Number (All)	97.1	96.5	97.5	100.0
Singular	97.4	95.8	98.3	100.0
Plural	96.8	97.1	96.6	100.0

Table 44. Accuracy rates for gender by group

	Intermediate	Advanced	Native	Totals
Accurate	114	110	134	358
Inaccurate	25	3	1	29
Totals	139	113	135	387

Table 45. Accuracy rates for number by group

	Intermediate	Advanced	Native	Totals
Accurate	232	137	158	527
Inaccurate	6	5	0	11
Totals	238	142	158	538

L2 subjects were more accurate in the comprehension task on items testing for number (97 percent) than for gender (89 percent), as was found for the elicited production experiment. The accuracy rate for gender agreement is somewhat higher in the comprehension task than the elicited production task, where the overall accuracy rates were 83 percent for gender overall and 86 percent for known gender. For intermediate subjects, however, the accuracy rates are essentially equal across comprehension and elicited production: between 81 and 83 percent for gender, and 98 percent for number. Advanced subjects' gender accuracy is higher in the comprehension task than the production task (97 percent for comprehension, 87 percent for production), but number accuracy is almost identical in the comprehension and production tasks (97 percent for comprehension, 99 percent for production), though near ceiling across both tasks.

Comprehension by feature

Gender

The contingency between feature and accuracy is significant overall, with L2 subjects significantly more accurate in the interpretation of feminine clitics than masculine ones ($X^2 = 8.92$, df = 1, $p<.005$; see Table 46). In other words, subjects are more likely to incorrectly interpret a masculine clitic as referring to a feminine object than the reverse, indicating the use of masculine clitics as a default in comprehension. A contingency is found for intermediate subjects ($X^2 = 9.71$, df = 1, $p<.002$; see Table 48), but advanced subjects show too few errors to test statistically (see Table 47). Due to the extremely high rate of accuracy of the advanced group, I cannot speculate on whether there is might be an interaction between proficiency level and error type.

Table 46. Comprehension of gender for all L2 subjects:

Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	109	115	224
Inaccurate	22	6	28
Totals	131	121	252

Table 47. Comprehension of gender for advanced L2 subjects:

Feature versus accuracy.

	Masculine	Feminine	Totals
Accurate	57	53	110
Inaccurate	2	1	3
Totals	59	54	113

Table 48. Comprehension of gender for intermediate L2 subjects:

Feature versus accuracy

	Masculine	Feminine	Totals
Accurate	52	62	114
Inaccurate	20	5	25
Totals	72	67	139

Number

Of the 11 errors in the comprehension of number in clitics, 5 involved singular clitics and 6 plural. There is no contingency between number feature and accuracy for L2 subjects ($X^2 = .75$, df = 1, $p<.2$; see Table 49). Of the 11 errors, 5 were made by advanced subjects and 6 by intermediate subjects; due to the low occurrence of errors in number, I will not provide a breakdown by group.

Table 49. Comprehension of number for all L2 subjects:

Feature versus accuracy

	Singular	Plural	Totals
Accurate	186	183	354
Inaccurate	5	6	10
Totals	183	181	364

5.4.3 Discussion

For the variable of gender, the results of this study show that default inflection surfaces in comprehension; furthermore, the same (masculine) default is chosen across comprehension and production. This runs counter to the suggestion that morphological variability is strictly a production-based phenomenon, and that the high processing demands on production are to blame for morphological variability. If these pressures were the sole cause of variability, removing them should remove variability as well. As I have shown here, removing production pressures does not eliminate variability. This is not to say that processing demands in production do not play a role in explaining the variability found in Sections 5.2 and 5.3, as well as other studies that have found variability in production (Lardiere 1998a,b, Prévost and White 2000b, Bruhn de Garavito 2003a,b, among others). Production pressures might explain why production lags behind comprehension; however, they cannot be invoked as the only explanation for the nature of morphological variability. In fact, it is not at all clear that production lags behind comprehension in these experiments: this pattern holds only for the advanced group's performance for gender. The intermediate group performed at essentially the same level across production and comprehension for both gender and number.

The predictions regarding the use default inflection were borne out for gender, in that masculine inflection acted as a default in feminine contexts. The result for number, however, is not consistent with predictions, as no feature asymmetry was found for this category in comprehension. With an overall accuracy rate of 97 percent, however, L2ers generally do not have a problem with the comprehension of number. Most likely, these number errors are just noise in the data.

Regarding the performance of the proficiency groups, L2 speakers show no difference from the NS group on feminine clitics. When it comes to interpreting feminine clitics, L2 speakers at all levels interpret these as unambiguously feminine, as do the NSs. However, masculine clitics are not interpreted with the same level of accuracy as was found for natives; L2 speakers, on the whole, differ from NSs in that they are willing to accept masculine clitics in feminine contexts. The problem with

comprehension, therefore, seems to be in L2ers' willingness to overextend masculine clitics. Only the advanced group, when considered separately, shows native-like performance across the board; this group, unlike the intermediate group, recognizes that *lo* and *los* are only acceptable when referring to masculine NPs.

One potential complicating factor for the pair of experiments reported in Sections 5.3 and 5.4 concerns the criteria for establishing knowledge of gender. Here, the consistent and correct use of a determiner and/or the correct naming of gender on the vocabulary task was taken as evidence of knowledge of gender. However, in Section 5.2, masculine gender surfaced as a default in the spontaneous production of determiners. Potentially, masculine gender could have surfaced as a default in the determiners that were taken as evidence for knowledge of gender in the elicited production experiment and the vocabulary naming task. The criteria for showing knowledge of gender may in fact be biased: subjects may demonstrate "knowledge" of masculine gender when, in fact, they do not know the gender of a masculine item.[51] Furthermore, with only two options to choose from (masculine or feminine), subjects can haphazardly guess the gender of an NP in a vocabulary naming task or while producing the NP with a determiner, and in many cases get it right. This is a problem that is impossible to circumvent, since knowledge of gender cannot be directly accessed.

5.5 Conclusion

The results of the three experiments described in this chapter support the main hypothesis of this dissertation. Across all three experiments, systematic patterns in errors emerged; in all the syntactic categories under analysis—determiners, clitics,

[51] It is likely, given the use of masculine defaults in determiners, that masculine gender contains more "false positives" for demonstrating knowledge of gender than feminine gender does. We would therefore expect that the category of masculine "known gender" might be artificially inflated. Correspondingly, we would expect more errors in masculine known gender to surface. In reality, we see very few errors in masculine known gender, which might suggest that this category is not falsely inflated. It would be highly problematic if the feminine "known gender" category were inflated rather than the masculine category, since we would not be able to determine whether errors in feminine gender were due to underspecification (as predicted), or false identification of knowledge of gender. Thus, the problem of identifying knowledge of gender is a serious one, but it would be even more serious if learners were biased toward falsely identifying knowledge of feminine gender.

and adjectives— masculine inflection was the choice of default where errors occurred. Number showed considerably less variability than gender. Nevertheless, in the two production experiments, singular number occurred as a default in plural contexts. No default was found in comprehension for number, and the accuracy rate was around 97 percent overall, suggesting complete mastery of number.

The first experiment examined a corpus of spontaneous production data from 11 L2 speakers, and found that masculine determiners occurred in feminine contexts significantly more frequently than feminine determiners occurred in masculine contexts. From this, I concluded that masculine determiners are defaults in spontaneous production. Number errors turned out to be extremely rare in the corpus, but out of three number errors, all involved the substitution of singular inflection for plural inflection.

The second experiment examined elicited production data involving clitics and adjectives, and found that masculine inflection in both cases occurred in feminine contexts significantly more than the reverse, confirming the trend found for spontaneous production. The consistent use of correctly-inflected determiners served as a diagnostic for separating NPs into two categories: those that the subject appeared to know the gender of, and those that the subject did not appear to know the gender of. When only items of known gender were considered, the same systematic substitution of masculine for feminine inflection was found. Furthermore, in isolating only errors that involved known gender, the vast majority of errors in both clitics and adjectives involved the use of masculine gender in feminine contexts. Apparently, even when a subject knows that a NP is feminine, masculine gender surfaces as a default. From the persistent variability in gender usage in known contexts— specifically when feminine gender is at stake— it is not the case that correct encoding of gender automatically entails the correct use of agreement (as suggested in Carroll 1989). The suppliance of agreement in feminine contexts remains problematic despite the knowledge of feminine gender. When gender is known, masculine agreement appears to "come for free"; feminine agreement, however, does not.

The third experiment found that default inflection extends to the domain of comprehension. L2ers made significantly more errors in the interpretation of masculine clitics than in the interpretation of feminine clitics, meaning that masculine clitics *lo* and *los* were extended to refer to feminine NPs. Feminine clitics, meanwhile, showed essentially no variability, in that *la* and *las* were consistently interpreted as referring to feminine NPs. Comprehension of number proved to be unproblematic, for the most part. Comprehension did not exceed production accuracy for the intermediate subjects, though the advanced subjects did perform better in comprehension than production of gender.

When viewed alongside the results of the production studies, the results of the comprehension experiment show that comprehension and production are not qualitatively different in terms of the use of default inflection. In theoretical terms, I take this to mean that the same mechanism of lexical insertion—and the insertion of underspecified inflection into the fully-specified syntax as an account for error patterns—can and should be extended to comprehension.

The finding that default inflection extends to comprehension is relevant to the debate over whether variability is an issue of representation or computation. A representational deficit would place the source of variability in an inability to represent functional features: for Hawkins (1998, 2001; also Hawkins & Chan 1997) in particular, the problem lies in L2ers' inability to select for functional features not represented in the L1. A computational deficit, on the other hand, attributes variability to mapping problems between syntax and morphology (Lardiere 2000, 2005), or communication pressures (Prévost and White 2000b). Previously, I stated that communication pressures alone cannot explain variability, but they can explain why production lags behind comprehension in some studies (though not this one). Mapping, however, may be more successful in explaining across-the-board variability, as it places the source of the problem in the organization of features. Yet attributing variability to mapping alone cannot explain why some features act as defaults and others do not—why should masculine clitics map onto syntactically feminine terminal nodes, for example, and not the reverse?

I propose that the unidirectionality of 'mapping' problems is, in fact, a representational issue. I do not use the term deficit here because the feature representations of L2ers appear to be entirely native-like, and NSs, by definition, do not have deficits. Recall that in Chapter 2 I used data from (native) Spanish to establish that masculine and singular are underspecified and therefore act as defaults; throughout this dissertation these were the defaults adopted by L2ers. The representational issue appears to lie in the asymmetrical representation of features, which in turn leads to the use of default inflection. By attributing variability to representational issues, we can capture the fact that comprehension is a domain of underspecification: both comprehension and production operate with the same feature representations and underspecification applies in both. The difference between this stance on representational issues and Hawkins', for example, is that these representational issues lie strictly in the domain of (post-syntactic) morphology, rather than syntax.

In Chapter Two, I presented evidence to show that the production of inflection should not be adopted as the criterion for the acquisition of functional features/projections (following arguments from Lardiere 1998a,b). For the moment, let us put this argument aside and assume that production is a valid criterion. By adopting this criterion, we are led to believe that the L2ers had not acquired the feature of gender. The accuracy of the advanced L2 group is, for the most part, different from that of the NSs; this is true in the production of gender in clitics, gender and number in adjectives, and the comprehension of number. In fact, the only measure on which the advanced group's performance is *not* significantly different from that of the NSs is in the comprehension of gender. Under such a strict criterion of acquisition of morphology, these statistical differences between NSs and advanced L2ers would be taken to mean that the features of gender and number have not been successfully acquired. However, production in masculine contexts is consistently native-like across comprehension and production. This raises a theoretical question: Is it possible to have acquired masculine gender—as measured by near-perfect performance in comprehension and production— without having acquired larger

organizing feature of gender? I believe the answer is no, and I will explain why below.

First, as previously noted by White (in press), a lack of functional categories/projections might be expected to yield randomness in errors,[52] not the systematic adoption of one default form over another that we find in the three experiments reported here. Thus a missing feature of gender should not yield consistent use of masculine gender. Second, the substitution of masculine for feminine, for example, means that the two terms are somehow linked: if they were not, why would masculine surface instead of any other feature (be it tense, person, or number)?[53] The L2ers appear to have built up a representation for the feature of gender, since errors inevitably involve the substitution of one gender term for another. Something must be linking masculine and feminine together as a natural class in the L2 grammar.

Returning to the feature representations I discussed in Chapter 2, the pattern of systematic substitutions is consistent with a theory that posits the (hierarchical) organization of features into classes. The systematic substitution of masculine for feminine provides an argument for this type of representation, as the presence of an organizing feature (node) of gender explains why this substitution occurs. The L2ers have acquired gender—as a natural class, hence presumably as a feature.[54] Thus the existence of systematic substitutions rather than random ones argues for the successful acquisition of the feature of gender in L2 Spanish, despite the lack of

---

[52] An exception to this might be in the case of a zero vs. overt marking alternation. If there is no feature, the default would presumably be zero, not a random alternation between zero and the overtly marked alternate (since the presence of an affix entails the presence of a syntactic projection, as I argued in Chapter 2). For Spanish gender, the alternation is not such a case, as both genders have overt marking; therefore a missing feature should predict randomness (in the form of errors going in both directions). I discuss the issue of zeros and missing inflection in more detail in Chapter 2.

[53] The suggestion that clitics be inflected for tense initially sounds absurd. Yet if we assumed that features were haphazardly arranged in bundles, there is no reason to expect for clitics *not* to represent tense. The combinatorial possibilities of what can constitute a "bundle" is enormous. Restricting these possibilities is a major argument for the use of feature geometries (Harley & Ritter 2002).

[54] I assume that, in the realm of inflection, natural categories in morphology are translatable as functional categories in syntactic terms. If masculine and feminine make up the larger organizing feature of gender from the point of view of syntax, and if masculine and feminine make up the natural category of gender from the point of view of morphology, and syntax and morphology are really one continuous stream (as is proposed under non-lexicalist theories like DM), then we really are talking about the same thing.

grammatical gender in the subjects' L1, English. To conclude, morphological variability can be used as an argument in favor of the acquisition of functional categories, even though it has often been used to as evidence of failure and deficits.

Chapter Six
Conclusion

This dissertation presented original research on variability in nominal and verbal morphology. In this concluding chapter, I will summarize the results of the experiments described in this thesis, and show how they support the underspecification-based hypothesis that I proposed in Chapter Three. Following this, I will raise some questions that this dissertation brings to light, and discuss some ways that we might go about answering them. I will conclude with some final remarks on what the study of L2 features might have in our understanding of linguistic competence.

6.1 Summary of major findings

This dissertation aims to explain an aspect of L2 morphological variability that has been largely unaddressed in the generative literature: the variants involved in default morphology. Previous research has focused on the syntactic consequences of morphological variability, putting aside the issues of what features are involved in variability, and what defaults learners resort to when variability arises. By looking at several different components of inflectional morphology across comprehension and production, it is clear that substitution errors are principled rather than random, and that the defaults that learners resort to are largely predictable; with these observations in mind, the phenomenon of morphological variability should not be dismissed as mere "performance".

In this dissertation, the discussion of features began with the adoption of underspecification and the establishment of feature inventories. Following insights in theoretical morphology, I assumed that features are represented asymmetrically, in that one feature within each opposed pair (e.g. masculine versus feminine) is underspecified. By assuming that underspecified features correspond to unmarked

values, the following features were diagnosed as being underspecified: masculine, singular, 3rd person, present, and nonfinite. Having established the set of features that are underspecified, I proposed the Morphological Underspecification Hypothesis (MUSH), which predicts that L2 errors are ones of underspecification, not feature clash. Morphological variability is, therefore, hypothesized to be governed by the representation of features.

The first set of experiments (reported in Chapter Four) presented original data on verbal morphology in L2 Spanish. The first experiment was a study of the spontaneous production of person, number, tense, and finiteness. L2ers used 3rd person, singular, present, and (to a lesser extent) nonfinite verbs as defaults in 1st/2nd person, plural, past, and finite contexts, respectively. The second experiment tested the comprehension of person and number in a written task; this task revealed essentially no variability, as accuracy was near 100 percent.

The second set of experiments (reported in Chapter Five) presented original data on nominal morphology in L2 Spanish: specifically, the features of gender and number. The first experiment was a study of the spontaneous production of gender and number in determiners. Masculine and singular defaults were attested in feminine and plural contexts, respectively. The second experiment was a study of the elicited production of gender and number in clitics and adjectives. Again, masculine and singular clitics and adjectives surfaced as defaults in feminine and plural contexts. Variability was attested even for advanced-proficiency speakers. The third experiment was a study of the comprehension of gender and number in clitics. This study demonstrated that morphological variability extends to comprehension. The observation of the same defaults in comprehension and production supports the claim that variability is governed by the representation of features. At the level of inflection, then, comprehension and production are not qualitatively different.

Taken together, the results support the claim that features are asymmetrically represented, and that default morphology—across comprehension and production--follows from these representations. Thus, the Morphological Underspecification Hypothesis is supported. In Chapter Two, I outlined the predictions of three

alternative hypotheses regarding morphological variability, and I now summarize how these hypotheses have fared in light of the data I have presented.

The Failed Functional Features Hypothesis (Hawkins & Chan 1997) correctly predicts variability across comprehension and production, but would seem to predict randomness rather than systematicity, and is therefore not supported. The Missing Surface Inflection Hypothesis (Prévost & White 2000b) predicts variability in production only; the occurrence of variability in comprehension is problematic for this proposal. Finally, Lardiere's Feature Assembly Hypothesis (Lardiere 2000, 2005) is supported, as it claims that features are responsible for variability and it would also seem to predict that variability could extend to comprehension. However, as I noted in Chapter Two, the hypothesis under its present formulation does not specify which features might be implicated and therefore does not allow us predict the asymmetrical relationships we have observed.

6.2 The representation of L2 features

One issue that arose in the discussion of the two experiments on the production of gender is the observation that, as L2ers get more advanced, the use of defaults gets more systematic to the point that errors of feature clash are totally unattested in some domains (e.g. known gender in the production of clitics). There is nothing in the MUSH that directly predicts this result. However, I would like to speculate about what might lie behind this effect.

In Chapter Two, I described two ways in which underspecification can be represented in the grammar: via feature bundle and via feature geometry. The feature-geometric approach offers a way to model the increasing systematicity of defaults. Following the logic of the Minimal Trees Hypothesis of Vainikka and Young-Scholten (1994, 1996) (see Chapter Three), we might assume that feature geometries are built up gradually. An intermediate-level L2er might have acquired the organizing node of gender, but still be in the process of building up the representation for the marked feature of feminine. An advanced-level L2er, on the other hand, would have

acquired the organizing node of gender, along with its dependant, feminine. These two representations are schematized in (1).

1. a) *Lower-proficiency L2er*

GENDER

|

b) *Higher-proficiency L2er*

GENDER

|

FEMININE

For an L2er with a representation like the one in (1a), feminine has not yet emerged, and so a feminine-for-masculine substitution error does not "clash" in the same way that it would for the more advanced L2er, who might have a representation like the one shown in (1b). As the L2er builds up his or her feature representations, we would expect for errors to become increasingly systematic.

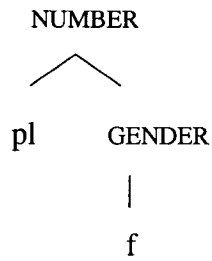
In addition to modeling the increasing systematicity of defaults, the use of a feature geometry captures the fact that masculine and feminine form a natural class of gender: substitution errors are therefore always contained within the feature category of gender. Therefore, as I argued in Chapter Five, the systematicity of substitution errors argues for the acquisition of "new" functional features like gender in L2 acquisition (and argues against the Failed Functional Features Hypothesis, which proposes that the acquisition of new uninterpretable features is impossible). Although I initially stated that either feature geometries or feature bundles might equally suffice so long as underspecification was represented (see Chapter Two), I believe that a feature-geometric approach holds a great deal of promise in its ability to capture and define the space within which variability operates.

A difficulty for the hypothesis proposed in this dissertation is the observation that certain predicted errors appear frequently (e.g. masculine for feminine, present for past), others appear occasionally (e.g. 3rd person for 1st/2nd person, singular for plural verbal morphology), others appear very rarely (e.g. singular for plural nominal morphology, nonfinite for finite) and still others do not appear at all (e.g. 1st singular for 1st plural, 3rd plural for 1st plural). Although we can successfully rule out (on the basis of feature clash) many of the errors that do not occur, there are also other

unattested errors that we have no way of ruling out. In this sense, the hypothesis overgenerates somewhat. I'll begin by discussing the representation of gender and number, and then turn to person and number.

The major generalization that we can draw about gender and number in nominal inflection, based on this dissertation and previous research (see Chapter Three and references therein), is that number is more accurate than gender. One possible explanation (along the lines of the FFFH) would attribute the difference between gender and number to the L1 feature inventory (gender is not present; number is present). However, this explanation does not go through, as previous research has shown that the existence of gender in the L1 does not guarantee L2 success (Bruhn de Garavito & White 2002) nor does the absence of gender in L1 guarantee L2 failure (as some subjects performed perfectly in some tasks). The organization of features offers an alternative explanation. We can model the higher rate of accuracy in number within a feature geometry by embedding gender within number as in (2), as has been previously proposed by Harley (1994):

2.



Under this representation, if we assume once again that representations are built up gradually, it would be impossible to represent gender without also representing number. This representation could then predict that gender is more likely to cause problems, as it is representationally more complex (in that it requires the presence of the number node above it). By assuming that structure is hierarchical and built up gradually, we could model the difference in accuracy between gender and number without having to resort to the prior existence of these features in the L1, an explanation that already seems to be problematic.

Constraining the predictions for verbal morphology is more complicated. The person/number representations that are most complex (1st person/2nd person plural) show no variability whatsoever, although as previously mentioned, this could be a by-product of the relative lack of 1st/2nd plural contexts. Further experimentation is required to determine whether errors in these contexts would emerge. This approach shows promise for verbal morphology when we compare the results obtained here, which show the use of 3rd person defaults in production, to the results reported by Mezzano (2003), who found that errors in person were somewhat random for learners at low levels of proficiency in L2 Spanish (see Chapter Four).

6.3 Questions for future research

In Section 6.2, I identified some possible directions for future research in the organization of features. In this section, I will outline some research goals at the level of methodology.

First, as I stated previously, the use of spontaneous production data in studying person and number is somewhat problematic, because L2ers do not use all possible person/number combinations with equal frequency. In particular, 1st and 2nd plural and 2nd singular forms were lacking. In the future, a task that elicits these contexts could control for this problem. In terms of comprehension, the overall accuracy in person and number agreement made it impossible to access informative data in comprehension. This problem could be rectified by testing participants at lower levels of proficiency, or perhaps participants who were naturalistic learners rather than classroom learners.

Second, for gender and number, I've suggested feature-geometric representations that appear to be plausible. If these structures are built up gradually, then the next step would be to test L2ers at the initial state. I have assumed that markedness determines feature specifications, but it would be worthwhile to consider whether L1 representations might play a role as well, as Lardiere (2000, 2005) suggests in the discussion of feature reassembly (see Chapter Three for a discussion).

Third, the analysis of errors is but one way to demonstrate asymmetry; another possibility would be through online tasks that tap into processing. Underspecified

features should behave differently than marked ones in the way that they are processed. A combination of methodologies would make the argument in favor of asymmetrical representations even more compelling.

Finally, I have attempted to provide an explanation for why errors are of a certain type. L2ers' feature representations are assumed to be asymmetrical and native-like; however, this fact alone does not explain why L2ers commit errors but NSs generally do not. The overuse of underspecified morphology is, under the hypothesis I propose, a product of a failure to block underspecified morphology— the grammar of L2ers might be seen to contain an ineffectively operating Elsewhere Principle (see Chapter Two). This proposal allows us to model variability in a straightforward way. Constraining the contexts in which we predict blocking to fail (and, correspondingly, variability to occur) is a question for future research.

6.4 Concluding remarks on (morphological) competence and performance

It is generally uncontroversial among generativists that syntax and phonology are structured; however, morphological representations are very often (implicitly) assumed to be unprincipled bundles of features. I hope to have contributed to the growing body of evidence that supports an important claim of morphologists: that morphological representations have organization and structure just as syntactic and phonological ones do, and that these morphological representations make up a part of linguistic competence. Ideally, the observations and arguments put forward here will be useful beyond the field of L2 acquisition; under the assumption that the interlanguage is, in fact, a language, these observations and arguments may inform the study of morphological features more generally. To take a specific example, L2ers' use of underspecified 3rd person and singular defaults is consistent with feature-geometric representations that are proposed to be universal (e.g. Harley & Ritter 2002).[55]

[55] The evidence for linguistic universals—in the form of feature representations that are consistent with UG—tells us that the L2ers have not built up 'rogue grammars', but this evidence does not provide an argument for the availability of UG in L2 acquisition. These representations could be built up based on

Throughout this dissertation, I have argued that variability is, to a large extent, a representational issue and therefore governed, by underlying linguistic competence. This discussion can be framed in terms of Lardiere's notion of morphological competence (see Chapter Three). However, I do not want to imply that morphological knowledge is necessarily a different kind of knowledge than syntactic or phonological knowledge. All of these areas are domains that consist of organized, structured representations, and we can aim to discover how the competence that contains these representations is acquired in the L2.

At the same time, I do not wish to totally dismiss the role of computational problems or "performance". Linguistic competence is accessible only indirectly (see White 2003:17), and performance may have a greater or lesser effect on the data we access depending on the task. Correspondingly, we may see more variability in one task as opposed to another. For the case of verbal morphology, relatively few errors in production were found; the total accuracy in comprehension, therefore, does not argue against the underspecification of features. Instead, it may reflect a very normal fact about (L1 and L2) acquisition, namely, that comprehension often exceeds production. Both production and comprehension are subject to "performance", and both domains offer insight into underlying competence.

To conclude, this dissertation aims to contribute to our understanding of the linguistic competence of the L2 learner in a new way. Many of the same questions that generative L2 research has asked about syntax can also be asked about morphology: What is the initial state of morphology? What is the role of markedness in representing morphological features? Are morphological representations always constrained by UG? What is the nature of near-native morphology? When we answer some of these questions, we may come to a better understanding of the place of morphology alongside the other levels of structure that have received more attention. By accepting that morphology is, in fact, organized and structured, perhaps we can begin to give it the attention it deserves.

---

evidence from the L2 through the course of acquisition; there is no need to invoke UG access to explain their feature representations.

References

Aguado-Orea, José Javier. 2004. *The acquisition of morpho-syntax in Spanish: Implications for current theories of development.* Unpublished PhD dissertation, University of Nottingham.

Barkin, Florence. 1980. The role of loanword assimilation in gender assignment. *Bilingual Review* 7(2): 105-112.

Battistella, Edwin L. 1990. *Markedness: The evaluative superstructure of language.* Albany, NY: SUNY Press.

Beard, Robert. 1995. *Lexeme-morpheme base morphology.* SUNY Albany Press, Albany.

Blom, Elma & Daniela Polisenska. 2006. The acquisition of inflectional morphology: Child L1 and child L2 Dutch. Poster presented at *Workshop on the Acquisition of Inflectional Morphology,* Budapest, May 27 2006.

Bobaljik, Jonathan D. 2006. Functional categories: Formal perspectives. Paper presented at *Generative Approaches to Language Acquisition North America (GALANA 2),* McGill University, Montreal, Canada, August 18 2006.

Bobaljik, Jonathan D. 2001. The implications of rich agreement: Why morphology does not drive syntax. In Karine Megerdoomian & Leora A. Bar-el, eds., *Proceedings of the 20th West Coast Conference on Formal Linguistics* (WCCFL 20), 82-95. Somerville, Mass.: Cascadilla Press.

Bonet, Eulàlia. 1995. Feature structure of romance clitics. *Natural Language and Linguistic Theory* 13: 607-647.

Brown, Roger. 1973. *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Bruhn de Garavito, Joyce. 2003a. The (dis)association between morphology and syntax: The case of L2 Spanish. In Silvina Montrul & Francisco Ordóñez, eds., *Linguistic theory and language development in Hispanic languages,* 398-417. Somerville MA: Cascadilla Press.

Bruhn de Garavito, Joyce. 2003b. Learners' competence may be more accurate than we think: Spanish L2 and agreement morphology. In Juana M. Liceras, Helmut Zobl & Helen Goodluck, eds., *Proceedings of the 6th Generative Approaches to Second Language Acquisition (GASLA 2006)*, 17-23. Somerville, MA: Cascadilla Proceedings Project.

Bruhn de Garavito, Joyce & Lydia White. 2002. L2 acquisition of Spanish DPs: The status of grammatical features. In Ana Teresa Perez-Leroux & Juana M. Liceras, eds., *The acquisition of Spanish morphosyntax: The L1/L2 connection*, 151-176. Dordrecht: Kluwer.

Carstairs-McCarthy, Andrew. 1998. How lexical semantics constrains inflectional allomorphy. In Geert Booij & Jap van Marle, eds., *Yearbook of morphology 1997*, 1-24. Dordrecht: Kluwer.

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Clahsen, Harald. 1988. Parameterized grammatical theory and language acquisition: A study of the acquisition of verb placement and inflection by children and adults. In Suzanne Flynn & Wayne O'Neil, eds., *Linguistic theory in second language acquisition,* 47-75. Dordrecht: Kluwer.

Clahsen, Harald & Pieter Muysken. 1986. The availability of Universal Grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research* 2: 93-119.

Clahsen, Harald & Pieter Muysken. 1989. The UG paradox in L2 acquisition. *Second Language Research* 5: 1-29.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2:225-252.

Conradie, Simone. 2005. *Verb movement parameters in Afrikaans: Investigating the Full Transfer Full Access hypothesis.* Unpublished PhD dissertation, McGill University.

Cowper, Elizabeth. 2004. The geometry of interpretable features: INFL in English and Spanish. Unpublished ms., University of Toronto.

Dulay, Heidi & Marina Burt. 1974. Natural sequences in child second language acquisition. *Language Learning* 24(1):37-53.

Español-Echevarría, Manuel & Philippe Prévost. 2004. On the L2 acquisition of number on Spanish quantifiers. In Alejna Brugos, Linnea Micciulla & Christine E. Smith, *Proceedings of the 28th Annual Boston University Conference on Language Development (BUCLD),* 135-64. Somerville, MA: Cascadilla Press.

Forcheimer, Paul. 1953. *The category of person in language.* Berlin: de Gruyter.

Franceschina, Florencia. 2001. Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research* 17(3): 213-247.

Franceschina, Florencia. 2002. Case and φ-feature agreement in advanced L2 Spanish grammars. In Susan H. Foster-Cohen, Tanja Ruthenberg & Marie Louise Poschen, eds., *EUROSLA Yearbook 2*, 71-86. Amsterdam: John Benjamins.

Goad, Heather, Lydia White & Jeff Steele. 2003. Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics* 48(3/4), 243-263.

Greenberg, Joseph H. 1966. *Language universals with special reference to feature hierarchies.* The Hague: Mouton.

Grimshaw, Jane. 1994. Minimal projection and clause structure. In Lust, B., Suñer, M., and Whiteman, J., eds., *Syntactic theory and first language acquisition: Cross-linguistic perspectives,* Volume 1. Hillsdale, NJ: Lawrence Erlbaum.

Halle, Morris. 1997. Distributed morphology: Impoverishment and fission. In Benjamin Bruening, Yoonjung Kang & Martha McGinnis, eds., *MITWPL 30: Papers at the interface.* MITWPL, Cambridge, 425-449.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Ken Hale & Samuel Jay Keyser, eds., *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111-176. Cambridge, MA: MIT Press.

Harley, Heidi. 1994. Hug a tree: Deriving the morphosyntactic feature geometry. In Andrew Carnie, Heidi Harley, & Tony Bures, eds., *MIT Working Papers in Linguistics 21:* 289-320.

Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78: 482-526.

Harris, James W. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22: 27-62.

Hawkins, Roger. 1998. The inaccessibility of formal features of functional categories in second language acquisition. Paper presented at the *Pacific Second Language Research Forum* (PacSLRF), Tokyo, March 1998.

Hawkins, Roger. 2001. *Second language syntax: a generative introduction.* Oxford: Blackwell.

Hawkins, Roger & Cecilia Yuet-hung Chan. 1997. The partial availability of universal grammar in second language acquisition: The 'Failed Functional Features Hypothesis'. *Second Language Research* 13(3): 187-226.

Hawkins, Roger & Florencia Franceschina. 2004. Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In Philippe Prévost & Johanne Paradis, eds., *The acquisition of French in different contexts*, 175–205. Amsterdam: John Benjamins.

Haznedar, Belma & Bonnie Schwartz. 1997. Are there optional infinitives in child L2 acquisition? In Elizabeth Hughes, Mary Hughes, & Annabel Greenhill, eds.,

Proceedings of the 21st Annual Boston University Conference on Language Development (BUCLD), 293-306. Somerville, MA: Cascadilla Press.

Herschensohn, Julia. 2000. *The second time round: Minimalism and L2 acquisition.* Amsterdam: John Benjamins.

Hyams, Nina. 1996. The underspecification of functional categories in early grammar. In Harald Clahsen, ed., *Generative perspectives on language acquisition: Empirical findings, theoretical considerations, crosslinguistic comparisons,* 99-127. Amsterdam: John Benjamins.

Jakobson, Roman. 1957/1971. Shifters, verbal categories and the Russian verb. *Selected writings, Vol. 2: Word and language*, 130-147. The Hague: Mouton.

Labov, William. 1972. *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Lardiere, Donna. 1998a. Case and tense in the 'fossilized' steady state. *Second Language Research* 14(1): 1-26.

Lardiere, Donna. 1998b. Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research* 14(4): 359-375.

Lardiere, Donna. 2000. Mapping features to forms in second language acquisition. In John Archibald, ed., *Second language acquisition and linguistic theory*, 102-129. Oxford: Blackwell.

Lardiere, Donna. 2005. On morphological competence. In Laurent Dekydtspotter, Rex A. Sprouse, & Audrey Liljestrand, eds., *Proceedings of the 7th Generative Approaches to Second Language Acquisition (GASLA 2004),* 178-192. Somerville, MA: Cascadilla Proceedings Project.

McCarthy, Corrine. 2006. What's missing from missing inflection?: Features in L2 Spanish. In Eva Dobler & Yukio Furukawa, eds., *McGill Working Papers in Linguistics* 20(2).

Mezzano, Gina. 2003. The development of Spanish verbal inflection in early stages of L2 acquisition. Undergraduate honors thesis, University of Illinois at Urbana-Champaign.

Meisel, Jurgen. 1991. Principles of Universal Grammar and strategies of language learning: Some similarities and differences between first and second language acquisition. In Lynn Eubank, ed., *Point counterpoint: Universal Grammar in the second language*, 324-76. Amsterdam: John Benjamins.

Montrul, Silvina. 1998. The L2 acquisition of dative experiencer subjects. *Second Language Research* 14(1), 27-61.

Montrul, Silvina A. 2004. *The acquisition of Spanish: Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition.* Amsterdam: John Benjamins.

Montrul, Silvina & Slabakova, Roumyana. 2002. The acquisition of morphosyntactic and semantic properties of the aspectual tenses preterite and imperfect. In Ana Teresa Pérez-Leroux & Juana Liceras, eds., *The acquisition of Spanish morphosyntax. The L1/L2 connection,* 131-149. Dordrecht: Kluwer.

Müller, Natascha. 1998. UG access without parameter setting: A longitudinal study of (L1 Italian) German as a second language. In Maria-Luise Beck, ed., *Morphology and its interfaces in second language knowledge*, 115-164. Amsterdam: John Benjamins.

Noyer, Rolf. 1992. Features, projections and affixes in autonomous morphological structure. Ph.D. dissertation, MIT.

Parodi, Teresa, Bonnie Schwartz & Harald Clahsen. 1997. On the L2 acquisition of the morphosyntax of German nominals. *Essex Research Reports in Linguistics* 15, 1-44.

Poeppel, David & Kenneth Wexler. 1993. The full competence hypothesis of clause structure in early German. *Language* 69(1): 1-33.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365-424.

Prévost, Philippe & Lydia White. 2000a. Accounting for morphological variation in L2 acquisition : Truncation or missing inflection ? In Marc-Ariel Friedemann & Luigi Rizzi, eds., *The acquisition of syntax,* 202-235. London: Longman.

Prévost, Philippe & Lydia White. 2000b. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16(2), 103-133.

Rohrbacher, Bernhard W. 1999. *Morphology-driven syntax: A theory of V to I raising and pro-drop.* Amsterdam: John Benjamins.

Sabourin, Laura. 2003. *Grammatical gender and second language processing: An ERP study.* PhD dissertation, University of Groningen.

Sagey, Elizabeth. 1986. *The representation of features and relations in non-linear phonology.* PhD dissertation, MIT.

Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12, 40-72.

Torres, Jenna. 2003. The Acquisition of Clitics in L2 Spanish: A Discourse/ Functional Perspective of Lexical NPs vs. Clitics. In Silvina Montrul & Francisco Ordóñez, eds., *Linguistic theory and language development in Hispanic languages: Papers from the 5th Hispanic Linguistics Symposium and the 4th Conference on the Acquisition of Spanish and Portuguese,* 447-467. Somerville, MA: Cascadilla Press.

Vainikka, Anne & Martha Young-Scholten. 1994. Direct access to X'-theory: Evidence from Korean and Turkish adults learning German. In Teun Hoekstra & Bonnie D. Schwartz, eds., *Language acquisition studies in generative grammar,* 265-316. Amsterdam: John Benjamins.

Vainikka, Anne & Martha Young-Scholten. 1996. Gradual development of L2 phrase structure. *Second Language Research* 12, 7-39.

Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages.* Oxford: Oxford University Press.

White, Lydia. 2003. *Second language acquisition and Universal Grammar: from initial to steady state.* Cambridge: Cambridge University Press.

White, Lydia. in press. Some puzzling features of L2 features. In Juana Liceras, Helmut Zobl & Helen Goodluck, eds., *The role of features in second language acquisition.* Mahwah, NJ: Lawrence Erlbaum.

White, Lydia, Elena Valenzuela, Martyna Kozlowska-Macgregor, & Yan-kit Ingrid Leung. 2004. Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25: 105-133.

Wunderlich, Dieter. 1996. Minimalist Morphology: The role of paradigms. In Geert Booij & Jap van Marle, eds., *Yearbook of morphology 1995,* 93-114. Dordrecht: Kluwer.

Appendix A.

Individual results: Person in spontaneous production (lexical verbs only)

	non-3rd		3rd	
Subject	**Errors**	**Correct**	**Errors**	**Correct**
Samantha (Int)	2	70	0	47
Beth (Int)	4	41	1	40
David (Int)	1	17	0	14
Christine (Int)	0	27	0	12
Sheila (Int)	3	31	1	38
Martha (Int)	0	28	0	21
Steve (Adv)	2	40	0	27
Annie (Adv)	1	55	0	58
Linda (Adv)	1	66	0	34
Rachel (Adv)	0	38	0	24
Tom (Adv)	0	38	0	23
Totals	**14**	**451**	**2**	**338**

Appendix B.

Individual results: Number in spontaneous production (verbal morphology)

	Singular: Lexical		Plural: Lexical		Singular: Cop/Aux		Plural: Cop/Aux	
Subject	**Errors**	**Corr.**	**Errors**	**Corr.**	**Errors**	**Corr.**	**Errors**	**Corr.**
Samantha (Int)	0	107	0	12	0	80	0	4
Beth (Int)	0	81	3	3	0	52	1	5
David (Int)	0	31	0	1	0	5	1	0
Christine (Int)	0	33	0	6	0	32	0	5
Sheila (Int)	0	36	0	37	0	77	0	32
Martha (Int)	0	41	4	4	0	8	3	1
Steve (Adv)	0	41	2	26	0	43	0	3
Annie (Adv)	1	93	0	20	0	52	0	12
Linda (Adv)	0	76	6	19	0	67	1	12
Rachel (Adv)	1	50	2	9	0	28	1	3
Tom (Adv)	0	45	1	15	0	41	0	8
Totals	**2**	**634**	**18**	**152**	**0**	**485**	**7**	**85**

Appendix C.

Individual Results: Past tense in spontaneous production

	Lexical verbs		Copular/auxiliary	
Subject	**Errors**	**Correct**	**Errors**	**Correct**
Samantha (Int)	0	17	0	11
Beth (Int)	1	6	1	3
David (Int)	1	17	0	0
Christine (Int)	1	24	0	10
Sheila (Int)	3	27	3	46
Martha (Int)	0	11	0	3
Steve (Adv)	1	21	1	10
Annie (Adv)	0	55	0	19
Linda (Adv)	1	51	0	23
Rachel (Adv)	3	15	0	2
Tom (Adv)	1	18	0	1
Totals	**12**	**262**	**5**	**131**

Appendix D.

Individual results: Gender in spontaneous production

	Masculine		Feminine	
Subject	**Errors**	**Correct**	**Errors**	**Correct**
Samantha (Int)	1	35	13	23
Beth (Int)	1	73	22	16
David (Int)	3	12	5	9
Christine (Int)	3	46	5	38
Sheila (Int)	1	79	18	48
Martha (Int)	1	16	3	10
Steve (Adv)	1	47	3	16
Annie (Adv)	0	58	2	45
Linda (Adv)	1	51	12	52
Rachel (Adv)	1	29	4	19
Tom (Adv)	0	22	2	21
Totals	**13**	**468**	**89**	**297**

Appendix E.

Individual Results: Gender in elicited production of clitics

	Masculine		Feminine		Masculine Known		Feminine Known	
Subject	**Errors**	**Corr.**	**Errors**	**Corr.**	**Errors**	**Corr.**	**Errors**	**Corr.**
1	0	10	0	11	0	8	0	9
15	0	11	0	9	0	8	0	9
16	1	4	1	5	0	2	1	3
3	0	7	5	7	0	6	4	6
4	3	10	0	9	0	6	0	6
11	0	11	0	11	0	8	0	9
33	0	7	3	7	0	6	2	5
27	2	11	6	7	0	9	4	4
20	0	1	0	1	0	1	0	1
Advanced: Totals	**6**	**72**	**15**	**67**	**0**	**54**	**11**	**52**
6	0	9	0	10	0	7	0	7
35	0	8	1	5	0	7	1	4
17	1	10	3	5	0	8	1	5
22	2	5	2	11	2	5	1	7
24	1	7	0	6	1	4	0	5
21	0	8	0	10	0	5	0	7
34	0	5	2	5	0	5	1	3
31	1	10	5	5	0	8	3	2

Individual Results: Gender in elicited production of clitics
(Continued from previous page)

30	0	15	11	3	0	14	10	3
29	0	10	6	5	0	6	3	3
28	0	2	3	3	0	1	2	0
26	1	6	1	5	1	6	0	0
19	0	2	0	0	0	1	0	0
32	1	2	0	0	0	1	0	0
25	3	6	0	8	2	3	0	5
Intermediate: Totals	**10**	**105**	**34**	**81**	**6**	**81**	**22**	**51**
Totals: All	**16**	**177**	**49**	**148**	**6**	**135**	**33**	**103**

Appendix F.

Individual Results: Gender in elicited production of adjectives

	Masculine		Feminine		Masculine Known		Feminine Known	
Subject	**Errors**	**Corr.**	**Errors**	**Corr.**	**Errors**	**Corr.**	**Errors**	**Corr.**
1	0	5	0	8	0	5	0	8
15	0	13	2	9	0	10	1	9
16	1	11	5	9	1	8	3	8
3	0	2	1	2	0	2	1	2
4	2	3	1	4	0	1	0	3
11	0	10	0	8	0	5	0	7
33	0	4	2	1	0	0	2	1
27	0	8	1	3	0	5	1	2
20	1	9	0	15	1	8	0	13
Advanced: Totals	**4**	**65**	**12**	**59**	**2**	**44**	**8**	**53**
6	0	2	3	1	0	1	1	1
35	0	7	2	2	0	6	2	1
17	0	13	5	14	0	12	4	12
22	0	10	3	6	0	7	1	5
24	1	4	4	7	1	3	2	6
21	0	8	8	1	0	4	7	1
34	2	5	2	6	1	3	0	5
31	0	8	5	2	0	5	2	1

Individual Results: Gender in elicited production of adjectives
(continued from previous page)

30	0	8	1	1	0	8	1	1
29	1	1	0	2	0	1	0	1
28	1	6	2	4	0	1	1	2
26	1	8	2	9	0	6	0	5
19	1	11	11	9	0	9	9	9
32	0	7	5	1	0	4	3	1
25	2	4	1	8	0	1	1	8
Intermediate: Totals	**9**	**102**	**54**	**73**	**2**	**71**	**34**	**59**
Totals: All	**13**	**167**	**66**	**132**	**4**	**115**	**42**	**112**

Appendix G.

Individual Results: Gender in Comprehension

	Masculine		Feminine	
Subject	**Errors**	**Correct**	**Errors**	**Correct**
1	0	8	0	7
15	0	7	0	6
16	0	4	0	4
3	0	8	1	6
4	2	4	0	4
11	0	8	0	7
33	0	2	0	6
27	0	8	0	6
20	0	8	0	7
Advanced: Totals	**2**	**57**	**1**	**53**
6	1	3	0	7
35	1	6	0	7
17	1	6	0	4
22	1	7	0	7
24	2	0	1	4
21	1	4	0	4
34	0	3	0	1
31	1	7	0	6
30	1	3	0	5

Individual Results: Gender in Comprehension
(Continued from previous page)

29	0	4	0	4
28	5	1	2	1
26	2	1	0	2
19	2	2	0	4
32	1	2	2	2
25	1	3	0	4
Intermediate: Totals	**20**	**52**	**5**	**62**
Totals: All	**22**	**109**	**6**	**115**

Appendix H.
Test materials: Comprehension of person and number

Instructions: Please circle the letter corresponding to the item that best completes the sentence.

Ejemplo: _____ están sucios.
a) el sillón b) los platos c) la camisa d) las cucharas e) el suéter

---

1. _____ dos lápices sobre la mesa. **(c); filler**
a) está b) ha c) hay d) estar e) soy

2. _____ viene a la fiesta esta noche. **(e); 3s irregular**
a) yo b) tú c) nosotros d) ellos e) María

3. _____ dibujas los árboles. **(b); 2s regular**
a) tú b) María c) ellas d) nosotros e) yo

4. _____ mal tiempo en Montreal. **(c); filler**
a) es b) está c) hace d) hay e) son

5. _____ vamos a la librería hoy. **(d); 1pl irregular**
a) María b) ellos c) yo d) nosotros e) tú

6. _____ pongo la mesa antes de cenar. **(c); 1s irregular**
a) ellos b) nosotros c) yo d) tú e) Enrique

7. _____ vivimos en Guadalajara. **(e); 1pl regular**
a) tú b) María c) yo d) ellos e) nosotros

8. _____ almuerzas en la cafetería. **(c); 2s irregular**
a) nosotros b) Enrique c) tú d) María y Enrique e) yo

9. _____ levanto a las ocho. **(a); filler**
a) me b) se c) está d) te e) es

10. _____ han estudiado por muchos años. **(d); filler**
a) María b) estamos c) la computadora d) los médicos e) nosotros

11. _____ limpia el baño. **(b); 3s regular**
a) nosotros b) Enrique c) tú d) ellos e) yo

12. _____ estoy feliz porque hace buen tiempo. **(a); 1s irregular**
a) yo b) tú c) María d) María y yo d) ellos

13. _____ se encuentran en una escuela. **(c); filler**
a) la nieve b) los gatos c) los alumnos d) los bosques e) el país

14. _____ habla con el veterinario. **(d); 3s regular**
a) tú b) yo c) María y Enrique d) Enrique e) Enrique y yo

15. _____ juega al tenis en el parque. **(b); 3s irregular**
a) nosotros b) María c) yo d) tú e) ellos

16. _____ pueden bailar bien. **(c); 3pl irregular**
a) tú b) Enrique c) ellos d) yo e) nosotros

17. _____ somos de Italia. **(d); 1pl irregular**
a) yo b) él c) ellas d) nosotros e) tú

18. _____ lloviendo. **(d); filler**
a) hace b) hay c) son d) está e) es

19. _____ son altos. **(a); filler**
a) María y Enrique b) ellas c) nosotros d) María e) las chicas

20. _____ vuelve a medianoche todos los días. **(a); 3s irregular**
a) María b) nosotros c) yo d) tú e) Enrique y María

21. _____ escribo cartas románticas. **(e); 1s regular**
a) nosotros b) tú c) Enrique d) Enrique y María e) yo

22. _____ llega al trabajo muy temprano. **(a); 3s regular**
a) María b) ellos c) yo d) Enrique y yo e) tú

23. _____ escrito un ensayo sobre la guerra. **(b); filler**
a) es b) he c) son d) iba e) soy

24. _____ siempre cierran la ventana para que no entren moscas. **(c); 3pl irregular**
a) nosotros b) Enrique c) ellos d) tú d) yo

25. _____ piensas que Lima es una ciudad peligrosa. **(b); 2s irregular**
a) yo b) tú c) él d) María y Juana e) nosotros

26. _____ duermes hasta las diez los fines de semana. **(a); 2s irregular**
a) tú b) nosotros c) ellos d) yo e) Enrique

27. _____ mienten a veces. **(c); 3pl irregular**
a) María b) Enrique y yo c) María y Enrique d) yo e) tú

28. _____ vamos al cine los fines de semana. **(d); 1pl irregular**
a) yo b) Enrique c) María y Enrique d) María y yo e) tú

29. _____ miro los pájaros en el parque. **(d); 1s regular**
a) María b) tú c) nosotros d) yo e) ellos

30. _____ lavas los platos con cuidado. **(e); 2s regular**
a) yo b) nosotros c) ellos d) ella e) tú

31. _____ sé hablar francés. **(d); 1s irregular**
a) tú b) Enrique y yo c) María d) yo e) ellas

32. _____ tomamos el autobús todos los días. **(a); 1pl regular**
a) nosotros b) tú c) yo d) María y Juana e) María

33. _____ comen las papas fritas. **(a); 3pl regular**
a) Enrique y María b) tú c) María d) nosotros e) yo

34. _____ llama Guillermo. **(d); filler**
a) son b) está c) su d) se e) es

35. _____ tocas el piano muy bien. **(c); 2s regular**
a) María b) yo c) tú d) nosotros e) ellos

36. _____ caminan en el bosque para ver los pájaros. **(a); 3pl regular**
a) María y Enrique b) Enrique c) Enrique y yo d) tú e) yo

37. _____ es alta. **(e); filler**
a) el señor b) Enrique c) yo d) tú e) la señora

38. _____ manejan demasiado rápido. **(c); 3pl regular**
a) Enrique b) tú c) ellos d) nosotros e) yo

39. _____ abro la ventana cuando hace calor. **(c); 1s regular**
a) tú b) ellos c) yo d) María e) nosotros

40. _____ arreglamos las bicicletas. **(b); 1pl regular**
a) ellos b) nosotros c) él d) tú e) yo

Appendix J.

List of vocabulary used in comprehension of gender and number experiment

masculine	*feminine*
anillo (ring)	bota (boot)
arete (earring)	blusa (blouse)
calcetines (socks)	bufanda (scarf)
cinturón (belt)	camisa (shirt)
cuaderno (notebook)	camiseta (t-shirt)
cuchillo (knife)	carta (letter)
espejo (mirror)	cartera (wallet)
guante (glove)	corbata (tie)
huevo (egg)	ensalada (salad)
libro (book)	falda (skirt)
paraguas (umbrella)	hamburguesa (hamburger)
pasaporte (passport)	llave (key)
periódico (newspaper)	maleta (suitcase)
reloj (watch)	papas (patatas) fritas (french fries)
tomate (tomato)	pelota (ball)
sandwich (sandwich)	pulsera (bracelet)
suéter (sweater)	raqueta (de tenis) (tennis racket)
tenedor (fork)	revista (magazine)
traje (suit)	silla (chair)
vaso (glass)	sopa (soup)
vestido (dress)	tarjeta postal (postcard)
zapato (shoe)	taza (cup)
vestido (dress)	toalla (towel)

Appendix K.

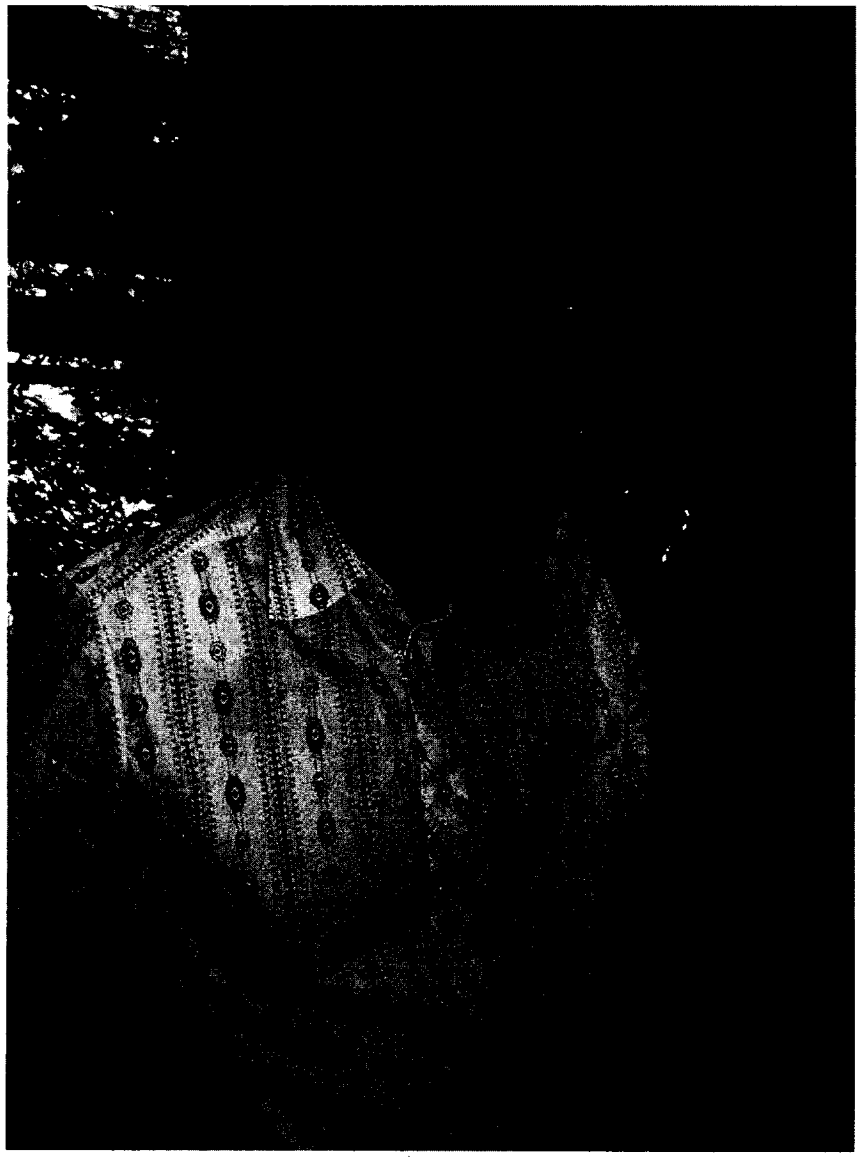
Sample pictures from elicited production of gender and number experiment

Dibujando hojas (drawing leaves-FEM)

Sample pictures from elicited production of gender and number experiment (continued from previous page)

Vendiendo libros (selling books-MASC)

Sample pictures from elicited production of gender and number experiment (continued from previous page)



Comiendo una manzana (eating an apple-FEM)

Appendix L.

Ethics Approval Certificates
McGill University and the University of Western Ontario
(see following two pages)